

Docket No.: 066662-0092

**PATENT**

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re Application of	:	Customer Number: 41552
	:	
Palsson, Bernhard	:	Confirmation Number: 1729
	:	
Application No.: 09/923,870	:	Group Art Unit: 1631
	:	
Filed: August 06, 2001	:	Examiner: Negin, Russell Scott
	:	
For: METHODS FOR IDENTIFYING DRUG TARGETS BASED ON GENOMIC SEQUENCE DATA	:	

**DECLARATION UNDER 37 C.F.R. § 1.132**

Mail Stop Amendment  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

I, Bernhard O. Palsson, declare as follows:

1. I am a Professor in the Department of Bioengineering at the University of California, San Diego (UCSD), San Diego, California. I am also an Adjunct Professor of Medicine at UCSD. I have held the former positions since joining UCSD in 1995, and the latter since 1998. I previously served on the faculty at the University of Michigan from 1984-1995.

2. I obtained a Bachelors of Science majoring in chemical engineering in 1979 from the University of Kansas, and earned my Ph.D. from the University of Wisconsin-Madison in Chemical Engineering in 1984. A copy of my curriculum vitae and a list of publications is attached as Exhibit 1.

3. I am an inventor or co-inventor on at least 35 U.S. Patents and a founder or co-founder of four life science companies. I co-founded Aastrom Biosciences, a public company that focuses on process technologies and devices for cell therapy applications, in 1989 and served as the Vice President of Developmental Research 1994-1995. I am the founder of

**EXHIBIT B**

Oncosis, a company focused on the purging of occult tumor cells in autologous bone marrow renamed Cyntellect and is focused on building research instrumentation for cell biology.

4. I also am a founder of Genomatica, Inc., which was incorporated in 1998 and is the exclusive licensee of the above-identified application. From 1998 to 2002 I served as the Company's Chief Executive Officer. Currently, I am the chair of Genomatica's Scientific Advisory Board.

5. I am named as the sole inventor on the above-identified patent application, U.S. patent application serial no. 09/923,870.

6. I have read the Office Action mailed December 18, 2008, and am very familiar with the prosecution history of this application. I understand that the claimed invention stands rejected for obviousness over the combination of references to Pramanik et al., *Biotech. and Bioengineering* 56:398-421 (1997) in view of Blattner et al., *Science* 277:1453-69 (1997) and in view of Kunst et al., *Rev. in Microbiol.* 142:905-12 (1991). The Examiner appears to equate the metabolic model described in Pramanik et al. with that described in the subject application and then combines it with two genome sequencing papers (Blattner et al. and Kunst et al. reporting the sequencing of *E. coli* and *B. subtilis* genomes, respectively) to conclude obviousness.

7. It is my opinion that the claimed invention is not obvious over Pramanik et al., Blattner et al. and Kunst et al. because (1) the use of genomic data would not have been expected to produce a predictable model, and (2) the opinions of respected colleagues in the field of metabolic modeling sincerely doubted that such a model would work, even after its publication.

8. One skilled in the art would not have been motivated to use genomic data to produce a stoichiometric matrix because the predictability of an *in silico* model relies on accuracy of the input data. Genome annotation is both incomplete and contains possible errors in the functional annotation of genes. Organism metabolism and reconstructing their metabolic networks from experimental data is scientifically very complex. At the time the invention was made, the metabolic network of a simple organism was known to have hundreds of reactions, substrates, products, co-factors and regulatory interventions and is dependent on many other events occurring within a cell. The predictability of a metabolic model depended on the

accuracy of this complex set of reactions and events. Therefore, metabolic models prior to my discovery employed either known kinetic or biochemical data of known reactions or events in order to ensure the most accurate model as possible.

9. It was incomprehensible to those in the field of metabolic modeling and metabolic engineering to incorporate reactions based on putative gene sequence homology without knowledge that the reaction existed in the modeled organism and without any known kinetic or biochemical information because incorrect deductions would decrease, rather than increase, the predictive capability of the resultant metabolic model. Therefore, those in the field would have discouraged incorporating putative metabolic reactions into an *in silico* model based on gene sequence homology to different organism because it could have led to incorporating erroneous reactions and a loss of a model's accuracy.

10. The unexpected nature of my discovery that genomic data can be incorporated into a model to construct a stoichiometric matrix without other knowledge is borne out by the disbelief of respected colleagues in the field at the time the invention was made. In addition to the concerns about the accuracy of the genomic data, my colleagues doubted that a metabolic reconstruction containing information about only 15% of the identified genes in *E. coli* and no regulatory information would be able to lead to meaningful computations about physiological functions.

11. Professor James E. Bailey, a respected and prestigious scientist in the field of metabolic engineering, is one such colleague who, on the one hand, found the success of my approach hard believe, but on the other hand if the model worked as purported, characterized my genomic-flux balance approach a "very major advance," a "breakthrough."

12. Professor Bailey passed away on May 9, 2001. At the time of my invention, Professor Bailey was a professor in the Institute of Biotechnology, ETH Zurich, Swiss Federal Institute of Technology in Zurich, Switzerland. Professor Bailey held a great deal of stature and respect in the field of metabolic engineering. Among Professor Bailey's many accomplishments include the foundational paper on Metabolic Engineering ('Towards a Science of Metabolic Engineering' Science 252:1688-1675, 1991, attached as Exhibit 2) and the publication of the standard text book in biochemical engineering (J.E. Bailey and D.F. Ollis, BIOCHEMICAL

ENGINEERING FUNDAMENTALS, 2<sup>nd</sup> edition 1986, 984 pages total, McGraw-Hill, New York). Publications in memory of Professor Bailey summarizing his accomplishments, scientific career and stature in the field of metabolic engineering are attached as Exhibits 3 and 4.

13. Attached as Exhibit 5 is a letter from Professor Bailey to me dated July 8, 1999, commenting on my publication describing the robustness of an *E. coli* metabolic model that used genomic data in its construction. A copy of the publication that precipitated Professor Bailey's letter is attached as Exhibit 6 (Edwards and Palsson, *Biotechnol. Prog.* 16:927-39 (2000).

14. With some tenacity Professor Bailey questions the operability of a genomic-flux balance model and the case studies reported in Exhibit 4, finding them "hard[] . . . to understand" (paragraph 2) and raising a number of questions regarding how the genomic model arrives at rates from stoichiometry (paragraph 3). Professor Bailey also questions the inclusion of 47 enzymes in addition to those implied by the genome sequences (paragraph 4). These steps were added based on biochemical information alone.

15. In the concluding paragraph Professor Bailey shows his very real skepticism and reservation by explicitly questioning whether the genomic model is correct or whether it works. Conversely, if the model operates as described Professor Bailey characterizes the discovery as a breakthrough. This skepticism and acknowledgement that the discovery is a breakthrough in the field of metabolic engineering are clear when Professor Bailey states:

This method, if it is correct and if it works, is a very major advance, because it give a formalism for determining growth rates and pathway rates without any knowledge of any kinetics. I must say it is a little hard to believe that this can be done, but maybe you have made a breakthrough. *Id.* (emphasis added).

As an opinion leader in the field, it is fair to say that Professor Bailey's skepticism characterized the attitudes of my colleagues at the time.

16. Since my initial publication of a genomic-flux balance metabolic model and the above interaction with Professor Bailey there have now been probably more than 100 subsequent publications from my laboratory and from the work of others in the field documenting the operability and predictive power of this type of metabolic model. This discovery was unexpected because we found that we could incorporate metabolic reactions based on genomic

deductions and the model was able to accurately predict metabolic capabilities of the modeled organism. As would be characterized by Professor Bailey now that there are several hundred subsequent publications to quell his initial skepticism, this discovery was a very major advance or breakthrough in the field. For example, for *E. coli* alone, about 70 such scientific studies were reviewed in June 2008 (Nature Biotechnology, 26: 659-667, 2008), attached as Exhibit 7. Notable predictions by the genomically-based reconstruction were the computation of the effects of gene knock-outs (Edwards and Palsson, *Proc. Nat'l Acad. Sci. USA*, 97:5528-33, (2000) (Exhibit 8); Covert et al., *Nature*, 429:92-96 (2004) (Exhibit 9) and the outcomes of adaptive evolution (Ibarra et al., *Nature*, 420:186-189 (2002) (Exhibit 10)). These predictions were unexpected and widely noticed by the scientific community.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements are made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that any such willful false statement may jeopardize the validity of the application or any patent issued thereon.

Banhe J Palsson

June 16<sup>th</sup> 2009

## CURRICULUM VITAE

Bernhard Ø. Palsson

Professor  
Department of Bioengineering  
Adjunct Professor  
Department of Medicine  
University of California-San Diego  
9500 Gilman Drive  
La Jolla, California 92093-0412

Date printed: June 18, 2009

### EDUCATION

- University of Iceland, Chemistry Lower Division, 1975-1977.
- University of Kansas, BS in Chemical Engineering, 1977-1979,  
Research Project: Boundary conditions for cell kinetic models
- University of Kansas, Graduate Courses in Biochemistry, 1978-1979.
- University of Wisconsin, Ph.D. in Chemical Engineering, 1979-1984,  
Dissertation: Mathematical Modeling of Dynamics and Control in Metabolic Networks. Minor:  
Mathematics and Biochemistry.

### PROFESSIONAL EXPERIENCE

- *Teaching assistant:*  
University of Wisconsin, 1980 - 1984 in process design and in process control.
- *Research assistant:*  
University of Iceland, 1976-1977,
  - a) fatty acid oxidizing enzymes in cardiac muscle and
  - b) statistical analysis of biochemical dataUniversity of Kansas, 1978-1979,
  - a) cell kinetics, and
  - b) diffusion in polymer matricesUniversity of Wisconsin, 1980-1984,
  - a) economic evaluation of fermentation products as feedstocks for the chemical industries,
  - b) development of non-linear control theory, and
  - c) mathematical modeling of dynamics and control in metabolic networks
- Assistant Professor, 1984-1990, Department of Chemical Engineering,  
The University of Michigan, Ann Arbor
- Co-Founder, 1989, Ann Arbor Stroma Inc (renamed Aastrom Biosciences Inc)
- Associate Professor, 1990-1995, Department of Chemical Engineering,  
The University of Michigan, Ann Arbor
- George Granger Brown Associate Professor of Chemical Engineering, 1991-1995, The  
University of Michigan, Ann Arbor
- Vice-President for Developmental Research, 1994-1995,  
Aastrom Biosciences, Inc., Ann Arbor
- Professor, 1995-present, Department of Bioengineering,  
University of California-San Diego
- Visiting Professor, 1996, Department of Process Biotechnology,  
The Technical University of Denmark, Lyngby
- Founder, Chairman, 1997-Present, CEO (Until 2002), Oncosis/Cyntellect,
- Co-Founder (1998), Chairman (Until 2002), CEO (Until 2000), Iceland Genomics Corporation

- Adjunct Professor of Medicine, UCSD, 1998-Present
- Co-Founder (1999), Chairman, Genomatica, 1999-2001
- Hougén Professor, University of Wisconsin, 2000.
- Galetti Professor of Bioengineering, UCSD, 2004-Present
- Faculty Member, Keio University, Tokyo, 2005-Present

## SCHOLARSHIPS, AWARDS, AND DISTINCTIONS

Graduation from the University of Kansas with highest distinction, 1979  
 Inclusion in Who's Who among students in American Colleges and Universities, 1979  
 Award for outstanding undergraduate research, 1979  
 Fellow of the Institute of International Education, 1977-1979  
 Fellow of Rotary International, 1979-1980  
 Experiment Station Research Assistantship in Chemical Engineering, 1982-1983  
 NATO Fellow, 1984  
 FIRST award from the NIH, 1987-1992  
 George Granger Brown Professorship, 1991-1995  
 Chemical Engineering Research Award, University of Michigan, 1993  
 Fellow of the American Institute for Medical and Biological Engineering, 1995  
 Fulbright Fellowship and Ib Henriksen Fellow, 1996  
 Hougén Professorship, University of Wisconsin, Madison, 2000.  
 Lindbergh-Carrel Prize, 2001  
 Bayer Lectures Honoree, 2001  
 FPB/AICHe Award, 2002  
 NIH Computational Approaches to Biological Systems Seminar Series, 2003  
 Keynote Address, ESCAT Meeting, Spain, May, 2003  
 Keynote Address, Mass Spectrometry in Systems Biology, Santa Fe, NM, February 2004  
 Plenary Talk, ISAC International Congress, Montpelier, France, May, 2004  
 Keynote Address, 5<sup>th</sup> International Conference on Systems Biology, Heidelberg, Germany, October, 2004.  
 Keynote Address, IBC Cell Line Development and Engineering Conference, St. Louis, November, 2005.  
 Election to National Academy of Engineering, February, 2006.  
 UCSD Chancellor's Associates Faculty Award for Excellence in research in Science and Engineering, 2006  
 Selected as one of 47 most influential labs for technology development over 10 years, *Nature Biotechnology*, March, 2006.  
 Plenary Talk, IECA 2006, Jeju, Korea, October, 2006.  
 Plenary Talk, Summit on Systems Biology, Richmond, Virginia, June 2007.  
 Plenary Talk, Metabolomics Society 3rd Annual Conference, Manchester, UK, June, 2007.  
 Richard S.H. Mah Lecture, Northwestern University, October, 2007.  
 Plenary Lecture, Metabolic Engineering VII, Puerto Vallarta, September, 2008.  
 Keynote Address, IBC Discovery 2 Diagnostics Conference, San Diego, October, 2008.  
 The Ernst W. Bertner Memorial Award, 2008.  
 Honorary Doctorate Degree, Chalmers University of Technology, May, 2009.

## PUBLICATIONS – Refereed

1. B. O. Palsson and K. J. Himmelstein (1981), "Determination of Steady-State Generation Time Distribution from Labeled Mitosis Experimental Data," *Cell and Tissue Kinetics*, **14**, 365-377.
2. B. O. Palsson, S. Fathi-Afshar, D. F. Rudd, and E. N. Lightfoot (1981), "Biomass as a Source of Chemical Feedstocks," *Science*, **213**, 513-517.
3. B. O. Palsson and E. N. Lightfoot (1984), "Mathematical Modeling of Dynamics and Control in Metabolic Networks. Part I. On Michaelis-Menten Kinetics," *J. theor. Biol.*, **111**, 273-302.

4. B. O. Palsson, R. Jamier, and E. N. Lightfoot (1984), "Mathematical Modeling of Dynamics and Control in Metabolic Networks. Part II. Simple Dimeric Enzymes," *J. theor. Biol.*, **111**, 303-321.
5. B. O. Palsson, H. Palsson, and E. N. Lightfoot (1985), "Mathematical Modeling of Dynamics and Control in Metabolic Networks. Part III. Linear Reaction Sequences," *J. theor. Biol.*, **113**, 231-259.
6. B. O. Palsson and E. N. Lightfoot (1985), "Mathematical Modeling of Dynamics and Control in Metabolic Networks. Part IV. Local Stability Analysis of Single Biochemical Control Loops," *J. theor. Biol.*, **113**, 261-277.
7. B. O. Palsson and E. N. Lightfoot (1985), "Mathematical Modeling of Dynamics and Control in Metabolic Networks. Part V. Static Bifurcations in Single Biochemical Control Loops," *J. theor. Biol.*, **113**, 279-298.
8. C. G. Economou, M. Morari, and B. O. Palsson (1986), "Internal Model Control: Extension to Non-linear Systems," *Ind. Eng. Chem. Process Des. Dev.*, **25**, 403-411.
9. B. O. Palsson and A. Joshi (1987), "On the Dynamic Order of Structured *Escherichia coli* Growth Models," *Biotechnology and Bioengineering*, **29**, 789-792.
10. B. O. Palsson (1987), "On the Dynamics of the Irreversible Michaelis-Menten Reaction Mechanism," *Chem. Eng. Sci.*, **42**, 447-458.
11. B. O. Palsson, A. Joshi, and S. S. Ozturk (1987), "Reducing Complexity in Metabolic Networks: Making Metabolic Meshes Manageable," *Fed. Proc.*, **46**, 2485-2489.
12. Joshi and B. O. Palsson (1987) "*Escherichia coli* Growth Dynamics; a Three Pool Biochemically Based Description," *Biotechnology and Bioengineering*, **31**, 102-116.
13. B.O. Palsson and T. Groshans (1988), "Mathematical Modeling of Dynamics and Control in Metabolic Networks. Part VI. Dynamic Bifurcations in Single Biochemical Control Loops," *J. theor. Biol.*, **131**, 43-53.
14. S. S. Ozturk, B. O. Palsson, and J. Dressman (1988) "Dissolution of Ionizable Drugs in Buffered and Unbuffered Solutions," *Pharm. Res.*, **5**, 272-282.
15. S. S. Ozturk, B. O. Palsson, B. Donohoe, and J. Dressman (1988) "Kinetics of Release from Enteric Coated Tablets," *Pharm. Res.*, **5**, 550-564.
16. S. S. Ozturk, B. O. Palsson, and J. Thiele (1989), "Control of Interspecies Electron Transfer Flow During Anaerobic Digestion: Dynamic Diffusion-Reaction Models for Hydrogen Gas Transfer in Microbial Flocs," *Biotechnology and Bioengineering*, **33**, 745-757.
17. G. M. Lee, T. K. Huard, and B. O. Palsson (1988), "Effect of Anchorage Dependency on Growth Rate and Monoclonal Antibody Production of Hybridoma Cells," *Biotech. Lett.*, **10**, 307-312.
18. J. M. Savinell, G. M. Lee, and B. O. Palsson (1989), "On the Orders of Magnitude of Epigenic Dynamics and Monoclonal Antibody Production," *Bioprocess Engineering*, **4**, 231-234.
19. G. M. Lee, T. K. Huard, M. S. Kaminski, and B. O. Palsson (1988), "Effect of Mechanical Agitation on Hybridoma Cell Growth," *Biotech. Lett.*, **10**, 625-628.
20. Ozturk, B. O. Palsson, A. R. Midgley, and C. R. Halberstadt (1989), "Transtubular Bioreactor: A Perfusion Device for Mammalian Cell Cultivation," *Biotechn. Techn.*, **3**, 55-60.
21. J. D. Keasling and B. O. Palsson (1989), "On the Kinetics of Plasmid Replication," *J. theor. Biol.*, **136**, 487-492.
22. G. M. Lee, T. K. Huard, and B. O. Palsson (1989), "Effect of Serum Concentration on Hybridoma Growth and Monoclonal Antibody Production at Various Initial Cell Densities," *Hybridoma*, **8**, 369-375.
23. S. S. Ozturk, M. E. Meyerhoff, and B. O. Palsson (1989), "Measurement of Ammonia and Glutamine in Cell Culture Media by Gas Sensing Electrodes," *Biotechn. Techn.*, **3**, 217-222.
24. Joshi and B. O. Palsson (1989), "Metabolic Dynamics in the Human Red Cell. Part I - A Comprehensive Kinetic Model," *J. theor. Biol.*, **141**, 515-528.
25. Joshi and B.O. Palsson (1989), "Metabolic Dynamics in the Human Red Cell. Part II - Interactions With The Environment," *J. theor. Biol.*, **141**, 529-545.
26. Joshi and B. O. Palsson (1990), "Metabolic Dynamics in the Human Red Cell. Part III - Metabolic Reaction Rates," *J theor. Biol.*, **142**, 41-68.
27. Joshi and B. O. Palsson (1990), "Metabolic Dynamics in the Human Red Cell. Part IV - Data Prediction and Some Model Computations," *J. theor. Biol.*, **142**, 69-85.



28. G. M. Lee, J. M. Savinell, and B. O. Palsson (1989), "Serum Can Act as a Shear Protecting Agent in Agitated Hybridoma Cell Cultures," *Hybridoma*, **8**, 639-645.
29. J. D. Keasling and B. O. Palsson (1989), "ColE1 Plasmid Replication: A Simple Kinetic Description from a Structured Model," *J. theor. Biol.*, **141**, 447-461.
30. B. O. Palsson, J. D. Keasling, and S. G. Emerson (1990), "The Regulatory Mechanisms of HIV Replication Predict Multiple Expression Rates," *Proc. Natl. Acad. Sci. (USA)*, **87**, 772-776.
31. S. S. Ozturk, J. B. Dressman, and B. O. Palsson (1989), "On the Use of the Quasi-equilibrium Assumption for Drug Dissolution," *Pharm. Res.*, **7**, 425-429.
32. S. S. Ozturk and B. O. Palsson (1990), "Effect of Initial Cell Density on Hybridoma Growth, Metabolism, and Monoclonal Antibody Production," *J. Biotech.*, **16**, 259-278.
33. S. S. Ozturk and B. O. Palsson (1990), "Loss of Antibody Productivity During Long-Term Cultivation of a Hybridoma Cell Line in Low Serum and Serum-Free Media," *Hybridoma*, **9**, 167-175.
34. S. S. Ozturk and B. O. Palsson (1990), "Chemical Decomposition of Glutamine in Cell Culture Media: Effect of Media Type, pH, and Serum Concentration," *Biotechnology Progress*, **6**, 121-128.
35. G. Ozturk, S. S. Ozturk, B. O. Palsson, T. A. Wheatley, and J. B. Dressman (1990), "Mechanism of Release from Pellets Coated with an Ethylcellulose-based Film," *Journal of Controlled Release*, **14**, 203-213.
36. G. M. Lee and B. O. Palsson (1990), "Immobilization Can Improve the Stability of Hybridoma Antibody Productivity in Serum-Free Media," *Biotechnology and Bioengineering*, **36**, 1049-1055.
37. I-der Lee and B. O. Palsson (1990), "A Comprehensive Model of Human Erythrocyte Metabolism: Extensions to Include pH Effects," *Biomed. Biochem. Acta*, **49**, 771-789.
38. S. S. Ozturk and B. O. Palsson (1991), "Physiological Changes During the Adaptation of Hybridoma Cells to Low Serum and Serum-free Media," *Biotechnology and Bioengineering*, **37**, 35-46.
39. G. M. Lee and B. O. Palsson (1990), "Simplified Methods of Making Alginate-Polylysine Microcapsules for Hybridoma Cell Culture Using RPMI 1640 Medium," *Biotechnology Techniques*, **4**(5), 341-344.
40. S. S. Ozturk and B. O. Palsson (1991), "Examination of Serum and Bovine Serum Albumin as Shear Protective Agents in Agitated Cultures of Hybridoma Cells," *Journal of Biotechnology*, **18**, 13-28.
41. S. S. Ozturk and B. O. Palsson (1990), "Effects of Dissolved Oxygen on Hybridoma Cell Growth, Metabolism and Antibody Production Kinetics in Continuous Culture," *Biotechnology Progress*, **6**, 437-446.
42. S. S. Ozturk and B. O. Palsson (1991), "Effect of Medium Osmolarity on Hybridoma Growth, Metabolism, and Antibody Production," *Biotechnology and Bioengineering*, **37**, 989-993.
43. Jerry Caldwell, Betty Locey, Michael F. Clarke, Stephen G. Emerson, and Bernhard O. Palsson (1991), "The Influence of Medium Exchange Schedules on Metabolic, Growth, and GM-CSF Secretion Rates of Genetically Engineered NIH-3T3 Cells," *Biotechnology Progress*, **7**, 1-8.
44. G. M. Lee, A. Varma, and B. O. Palsson (1991), "Application of Population Balance Model to the Loss of Hybridoma Antibody Productivity," *Biotechnology Progress*, **7**, 72-75.
45. S. G. Emerson, B. O. Palsson, and M. F. Clark (1991), "The Construction of High Efficiency Human Bone Marrow Tissue Ex Vivo," *J. Cell. Biochem.*, **45**, 268-272.
46. J. D. Keasling, B. O. Palsson, and S. Cooper (1991), "Cell-Cycle-Specific F Plasmid Replication: Regulation by Cell Size Control of Initiation," *J. Bacteriology*, **173**, 2673-2680.
47. J. Caldwell, B. O. Palsson, B. Locey, and S. G. Emerson (1991), "Culture Perfusion Schedules Influence the Metabolic Activity and Granulocyte-Macrophage Colony-stimulating Factor Production Rates of Human Bone Marrow Stroma Cells," *J. of Cellular Physiology*, **147**, 344-353.
48. R. M. Schwartz, B. O. Palsson, and S. G. Emerson (1991), "Rapid Medium Perfusion Rate Significantly Increases the Productivity and Longevity of Human Bone Marrow Cultures," *Proc. Natl. Acad. Sci. (USA)*, **88**, 6760-6764.

49. G. M. Lee and B. O. Palsson (1991), "Production of Monoclonal Antibody Using Free-suspended and Immobilized Hybridoma Cells: Effect of Serum," *Biotechnology and Bioengineering*, **38**, 821-830.
50. M. Javanmardian and B. O. Palsson (1991), "High-density Photoautotrophic Algal Cultures: Design, Construction, and Operation of a Novel Photobioreactor System," *Biotechnology and Bioengineering*, **38**, 1182-1189.
51. G. M. Lee, J. J. Gray, and B. O. Palsson (1991), "Effect of Trisodium Citrate Treatment on Hybridoma Cell Viability," *Biotechnology Techniques*, **5**(4), 295-298.
52. A.S.Y. Chuck and B. O. Palsson (1991), "The Population Balance Between Producing and Nonproducing Hybridoma Clones Is Very Sensitive to Serum Level, State of Inoculum, and Medium Composition," *Biotechnology and Bioengineering*, **39**, 354-360.
53. R. M. Schwartz, S. G. Emerson, M. F. Clarke, and B. O. Palsson (1991), "In Vitro Myelopoiesis Stimulated by Rapid Medium Exchange and Supplementation with Hematopoietic Growth Factors," *Blood*, **78**, 3155-3161.
54. M. Javanmardian and B. O. Palsson (1992), "Continuous Photoautotrophic Cultures of the Eukaryotic Alga *Chlorella vulgaris* Can Exhibit Stable Oscillatory Dynamics," *Biotechnology and Bioengineering*, **39**, 487-497.
55. Varma, F. Y. El-Awar, B. O. Palsson, S. G. Emerson, M. F. Clarke (1992), "Can Dexter Cultures Support Stem Cell Proliferation?" *J. Exp. Hematology*, **20**, 87-91.
56. J. M. Savinell and B. O. Palsson (1992), "Network Analysis of Intermediary Metabolism using Linear Optimization: I. Development of Mathematical Formalism," *J. theor. Biol.*, **154**, 421-454.
57. J. M. Savinell and B. O. Palsson (1992), "Network Analysis of Intermediary Metabolism using Linear Optimization: II. Interpretation of Hybridoma Cell Metabolism," *J. theor. Biol.*, **154**, 455-473.
58. S. S. Ozturk, M. R. Riley, and B. O. Palsson (1991), "Effects of Ammonia and Lactate on Hybridoma Growth, Metabolism, and Antibody Production," *Biotechnology and Bioengineering*, **39**, 418-431.
59. S. S. Ozturk and B. O. Palsson (1991), "Growth, Metabolic, and Antibody Production Kinetics of Hybridoma Cell Culture: I. Analysis of Data from Controlled Batch Reactors," *Biotechnology Progress*, **7**, 471-480.
60. S. S. Ozturk and B. O. Palsson (1991), "Growth, Metabolic, and Antibody Production Kinetics of Hybridoma Cell Culture: II. Effects of Serum Concentration, Dissolved Oxygen Concentration, and Medium pH in a Batch Reactor," *Biotechnology Progress*, **7**, 481-494.
61. J. M. Savinell and B. O. Palsson (1992), "Optimal Selection of Metabolic Fluxes for in vivo Measurement: I. Development of Mathematical Methods," *J. theor. Biol.*, **155**, 201-214.
62. J. M. Savinell and B. O. Palsson (1992), "Optimal Selection of Metabolic Fluxes for in vivo Measurement: II. Application to *E. Coli* and Hybridoma Cell Metabolism," *J. theor. Biol.*, **155**, 215-242.
63. G. M. Lee, M. S. Kaminski, and B. O. Palsson (1992), "Observations Consistent With Autocrine Stimulation of Hybridoma Cell Growth and Implications for Large-Scale Antibody Production," *Biotech. Letters*, **14**:4, 257-262.
64. J. D. Keasling, B. O. Palsson, and S. Cooper (1992), "Replication of the R6K Plasmid during the *Escherichia coli* Cell Cycle," *Journal of Bacteriology*, **174**:3, 1060-1062.
65. G. M. Lee, J. M. Savinell and B. O. Palsson (1992), "Characterization of Physiological Changes in 3H5/y2BA2 Hybridoma Cells During Adaptation to Low Serum Media," *J. of Microbiology and Biotechnology*, **2**(2), 141-151.
66. J. D. Keasling, B. O. Palsson, and S. Cooper (1992), "Replication of Prophage P1 is Cell Cycle-Specific," *J. Bacteriol.*, **174**:13, 4457-4462.
67. J. D. Keasling, B. O. Palsson, and S. Cooper (1992), "Replication of Mini-F Plasmids During the Bacterial Division Cycle," *Res. Microbiol.*, **143**, 541-548.
68. I-D. Lee and B. O. Palsson (1992), "A Macintosh Software Package for Simulation of Human Red Blood Cell Metabolism," *Computer Methods and Programs in Biomedicine*, **38**:4, 195-226.
69. G. M. Lee, B. K. Han, J. H. Kim, and B. O. Palsson (1992), "Effect of Calcium Chloride Treatment on Hybridoma Cell Viability and Growth," *Biotechnology Letters*, **14**, 891-896.

70. B.O. Palsson and I-D. Lee (1993), "Model Complexity has a Significant Effect on the Numerical Value and Interpretation of Metabolic Sensitivity Coefficients," *J. theor. Biol.*, **161**, 299-315.
71. R.M. Schwartz, J. Caldwell, M.F. Clarke, S.G. Emerson, and B.O. Palsson (1993), "The Influence of extra-cellular matrix and stroma remodeling on the productivity of long-term human bone marrow cultures," *Cytotechnology*, **10**, 217-224.
72. G. M. Lee, A. S. Chuck, and B. O. Palsson (1993), "Cell Culture Conditions Determine the Enhancement of Specific Monoclonal Antibody Productivity of Calcium Alginate-entrapped S3H5/y2bA2 Hybridoma Cells," *Biotechnology and Bioengineering*, **41**, 330-340.
73. Varma and B.O. Palsson (1993), "Metabolic Capabilities of *Escherichia coli*. I. Synthesis of biosynthetic precursors and cofactors," *J. theor. Biol.*, **165**, 477-502.
74. Varma and B.O. Palsson (1993), "Metabolic Capabilities of *Escherichia coli*. II. Optimal growth patterns," *J. theor. Biol.*, **165**, 503-522.
75. Varma, B.W. Boesch, and B.O. Palsson (1993), "Biochemical Production Capabilities of *Escherichia coli*," *Biotechnology and Bioengineering*, **42**, 59-73.
76. B.O. Palsson, S-H Paek, R.M. Schwartz, M. Palsson, G-M Lee, S. Silver and S.G. Emerson (1993), "Expansion of Human Bone Marrow Progenitor Cells in a High Cell Density Continuous Perfusion System," *BioTechnology*, **11**, 368-372.
77. S.E. Merritt and B.O. Palsson (1993), "Loss of Antibody Productivity is Highly Reproducible in Multiple Hybridoma Subclones," *Biotechnology and Bioengineering*, **42**, 247-250.
78. M.R. Koller and B.O. Palsson (1993), "Tissue Engineering: Reconstitution of Human Hematopoiesis ex vivo," *Biotechnology and Bioengineering*, **42**, 909-930.
79. G.M. Lee, A.S. Chuck, and B.O. Palsson (1993), "Effect of Fixation Temperature on Flow Cytometric Measurement of Intracellular Antibody Content of Hybridomas during Batch Culture," *Biotechnology Techniques*, **7**, 271.
80. M.E. Meyerhoff, M. Trojanowicz, and B.O. Palsson (1993), "Simultaneous Enzymatic/Electrochemical Determination of Glucose and L-Glutamine in Hybridoma Media by Flow-Injection Analysis," *Biotechnology and Bioengineering*, **41**, 964-969.
81. Varma, B.W. Boesch, and B.O. Palsson (1993), "Stoichiometric Interpretation of *Escherichia coli* Glucose Catabolism under Various Oxygenation Rates," *Applied and Environmental Microbiology*, **59:8**, 2465-2473.
82. M.R. Koller, S.G. Emerson, and B.O. Palsson (1993), "Large-Scale Expansion of Human Stem and Progenitor Cells from Bone Marrow Mononuclear Cells in Continuous Perfusion Cultures," *Blood*, **82**, 378-384.
83. G.M. Lee and B.O. Palsson (1993), "Stability of Antibody Productivity Is Improved When Hybridoma Cells Are Entrapped in Calcium Alginate Beads," *Biotechnology and Bioengineering*, **42**, 1131-1135.
84. B.O. Palsson, B.Q. Shen, M.E. Meyerhoff, and M. Trojanowicz (1993), "Simultaneous Determination of Ammonia Nitrogen and L-Glutamine in Bioreactor Media Using Flow Injection," *Analyst*, **118**, 1361-1365.
85. Varma and B.O. Palsson (1994), "Predictions for Oxygen Supply Control to Enhance Population Stability of Engineered Production Strains," *Biotechnology and Bioengineering*, **43**, 275-285.
86. M. Levee, G.M. Lee, S-H. Paek, and B.O. Palsson (1994), "Microencapsulated Human Bone Marrow Cultures: A Potential Culture System for the Clonal Outgrowth of Hematopoietic Progenitor Cells," *Biotechnology and Bioengineering*, **43**, 734-739.
87. D.J. Oh, M.R. Koller, and B.O. Palsson (1994), "Frequent Harvesting from Perfused Bone Marrow Cultures Results in Increased Overall Cell and Progenitor Expansion," *Biotechnology and Bioengineering*, **44**, 609-616.
88. C-G. Lee and B.O. Palsson (1994), "High-Density Algal Photobioreactors Using Light Emitting Diodes," *Biotechnology and Bioengineering*, **44**, 1161-1167.
89. G.M. Lee and B.O. Palsson (1994), "Monoclonal Antibody Production Using Free-Suspended and Entrapped Hybridoma Cells: A Comparative Study," *Biotechnology and Genetic Engineering Review*, **12**, 481-505.
90. A. Varma and B.O. Palsson (1994), "Stoichiometric Flux Balance Models Quantitatively Predict Growth and Metabolic By-Product Secretion in Wild Type *Escherichia coli* W3110, *Appl. Environ. Microbiol.*, **60:10**, 3724-3731.

91. G.M.Lee and B.O. Palsson (1994), "Enhanced Specific Antibody Productivity of Calcium Alginate-Entrapped Hybridoma is Cell Line Specific," *Cytotechnology*, **16**, 1-15.
92. Varma and B.O. Palsson (1995), "Parametric Sensitivity of Stoichiometric Flux Balance Models Applied to Wild Type *Escherichia coli* Metabolism," *Biotechnology & Bioengineering*, **45**, 69-79.
93. Varma and B.O. Palsson (1994), "Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use," *BioTechnology*, **12**, 994-998.
94. R.K. Mandalam and B.O. Palsson (1995), "*Chlorella vulgaris* (Chlorellaceae) Does Not Secrete Autoinhibitors at High Cell Densities," *American Journal of Botany*, **82**: 955-963.
95. C.G. Lee and B.O. Palsson (1995), "Light Emitting Diode-Based Algal Photobioreactor with External Gas Exchange," *J. of Fermentation and Bioengineering*, **79**:3, 257-263.
96. B.O. Palsson, D.J. Oh, and M.R. Koller (1995), "Replating of Bioreactor Expanded Human Bone Marrow Results in Extended Growth of Primitive and Mature Cells," *Cytotechnology*, **18**:125-131.
97. M.R. Koller, M.S. Bradley, B.O. Palsson (1995), "Growth Factor Consumption and Production in Perfusion Cultures of Human Bone Marrow Correlates with Specific Cell Production," *Experimental Hematology*, **23**:1275-1283.
98. M.R. Koller, I. Manchel, B.S. Newsom, M.A. Palsson, and B.O. Palsson (1995), "Bioreactor Expansion of Human Bone Marrow: Comparison of Unprocessed, Density-Separated, and CD-34-Enriched Cells," *J. Hematotherapy*, **4**:159-169.
99. C.A. Peng and B.O. Palsson (1995), "The Importance of Non-Homogeneous Concentration Distributions Near Walls in Tissue Engineering Bioreactors," special issue on Transport Phenomena, *Industrial & Engineering Chemistry Research*, **34**:3239-3245.
100. M.R. Koller, M.A. Palsson, I. Manchel, and B.O. Palsson (1995) "Long-Term Culture-Initiating Cell Expansion is Dependent of Frequent Medium Exchange Combined with Stromal and Other Accessory Cell Effects," *Blood*, **86**: 1784-1793
101. C.E. Rogers, M.S. Bradley, B.O. Palsson and M.R. Koller (1995), "Flow Cytometric Analysis of Human Bone Marrow Perfusion Cultures: Erythroid Development and Relationship with Burst-Forming Units-Erythroid (BFU-E)," *Experimental Hematology*, **24**: 597-604.
102. C.A. Peng, M.R. Koller, and B.O. Palsson (1996), "Unilineage Model of Hematopoiesis Predicts Self-Renewal of Stem and Progenitor Cells Based on Ex Vivo Growth Data," *Biotech & Bioeng.*, **51**:24-33.
103. JB Dressman, GM Derbin, G. Ismailos, C. Jarvis, A. Ozturk, BO Palsson, and TA Wheatley (1995), "Circumvention of pH dependent Release from Ethylcellulose-coated Tablets," *J. Controlled Release*, **36**: 251-260
104. Peng and B. O. Palsson, "Determination of Specific Oxygen Uptake Rates in Human Hematopoietic Cultures and Implications for Bioreactor Design," *Ann. Biomed. Eng.*, **24**: 373-381.
105. M.R. Koller, I. Manchel, M. A. Palsson, R.J. Maher and B. O. Palsson (1995), "Different Measures of Ex Vivo Hematopoiesis Culture Performance is Optimized Under Vastly Different Conditions," *Biotech & Bioeng.*, **50**:505-513
106. C.G. Lee and B.O. Palsson (1995), "Continuous Medium Perfusion Leads to Long-Term Cell Viability and Oxygen Production in High-Density Photobioreactors," *Biotech. Lett.*, **17**:1149-1154
107. C.A. Peng and B. O. Palsson, "Cell Growth and Differentiation on Feeder Layers is Predicted to be Influenced by Bioreactor Geometry," *Biotech & Bioeng.*, **50**:479-492
108. C.G. Lee and B.O. Palsson (1995), "Chlorella vulgaris Grown Under Red Light from Light Emitting Diodes has a Small Average Cell Size and Narrow DNA Distributions," *Biotech. Prog.*, **12**:249-256
109. Chuck and B.O. Palsson (1996), "Membrane Adsorption Characteristics Determine the Kinetics of Flow-Through Transfections" *Biotech & Bioeng.*, **51**:260-270
110. Chuck and B.O. Palsson (1996), "Consistent and High Rates of Gene Transfer Can be Obtained Using Flow-through Transduction Over a Wide Range of Retroviral Titers," *Human Gene Therapy*, **7**:743-750
111. P.G. Eipers, J.C. Krauss, B.O. Palsson, S.G. Emerson, R.F. Todd III, and M.F. Clarke (1995), "Retromediated Gene Transfer in Human Hematopoietic Cells Grown in Continuous Perfusion Culture Vessels," *Blood*, **86**:3754-3762

112. S. Andreadis and B.O. Palsson (1996), "Kinetics of Retrovirus Mediated Gene Transfer: The Importance of the Intracellular Half-life of Retroviruses," *J. Theoret. Biol.*, **182**: 1-20
113. M.R. Koller, M. Oxender, D.A. Brott and B.O. Palsson (1996), "flt-3 Ligand is More Potent Than c-kit Ligand for the Synergistic Stimulation of Ex Vivo Hematopoietic Cell Expansion," *J. Hematotherapy*, **5**: 449-459
114. Chuck, M.F. Clarke, and B.O. Palsson (1996), "Retroviral Infection is Limited by Brownian Motion," *Human Gene Therapy*, **7**:1527-1534
115. MR Koller, I Manchel, BA Brott and BO Palsson (1996), "Donor-to-Donor Variability in the Expansion Potential of Human Bone Marrow Cells is Reduced by Accessory Cells but not by Soluble Growth Factors," *Exp. Hematology*, **24**:1484-1493
116. S Andreadis, D Brott, OA Fuller, and BO Palsson (1997), "MMuLV-Derived Vectors Decay Intracellularly with a Half-life in the Range of 5.4 to 7.5 Hours," *Journal of Virology*, **71**:7541-7548
117. S Andreadis and BO Palsson (1997), "Coupled effects of Polybrene and Serum on Retroviral Stability are Coupled," *Human Gene Therapy*, **8**:285-291
118. BO Palsson and S Andreadis (1997), "The Physico-chemical Kinetic Factors that Govern Retrovirus-Mediated Gene Transfer," *Exp. Hematology*, **25**: 94-102
119. BQ Shen, MF Clarke and BO Palsson (1996), "Kinetics of Retrovirus Production from the ΨCRIP Producer Cell Line," *Cytotechnology*, **22**:185-195
120. MR Koller, I Manchel and BO Palsson (1997), "Importance of Parenchymal:stromal Ratio for Ex vivo Reconstitution of Human Hematopoiesis," *Stem Cells*, **15**:305-313
121. S Andreadis and BO Palsson (1997), "Cell cycle dependence of retroviral transduction: A matter of overlapping time scales," *Biotech & Bioeng.*, **58**:272-281
122. R.K. Mandalam and B.O. Palsson (1997), "Cell cycle of *Chlorella vulgaris* can deviate from the synchronous binary division model," *Biotech. Lett.*, **19**:587-591
123. K. Francis and B.O. Palsson (1997), "Effective Inter-cellular Communication Distances are Determined by the Relative Cytokine Secretion and Diffusion Time Constants," *Proc. Natl. Acad. Sci (USA)*, **94**: 12258-12262
124. JS Edwards and BO Palsson (1997), "How will Bioinformatics Influence Metabolic Engineering?," *Biotechnology & Bioengineering*, **58**:162-169
125. G.M. Lee, J. H. Choi, S.C. Jun and B.O. Palsson, (1998)"Temperature Significantly Affects Retroviral Vector Production and Deactivation Rates, and Thereby Determines Retroviral Titers," *Bioprocess Engineering*, **19**:343-349.
126. C.H. Schilling and B.O. Palsson (1998), "The underlying pathway structure of biochemical reaction networks," *Proc. Natl. Acad. Sci (USA)*, **95**: 4193-4198.
127. R.K. Mandalam and B.O. Palsson, "Elemental balancing of biomass and medium composition enhances growth capacity in high-density *Chlorella vulgaris* cultures," *Biotechnology and Bioengineering*, **59**:605-611 (1998).
128. S. Andreadis and Palsson, B.O. (1998), "Intracellular stability of retroviral vectors: mathematical modeling and kinetic experiments, *Cancer Research Therapy & Control* **7** (1-2): 77-85.
129. M.R. Koller, M.A. Palsson, I. Manchel, R.J. Maher, and B.O. Palsson (1998), "Tissue Culture Surface Characteristics Influence the Expansion of Human Bone Marrow Cells," *Biomaterials*, **19**: 1963-1972.
130. K. Francis, W. Holloway, R. Ramakrishna, and B.O. Palsson (1998), "Two New Pseudopod Morphologies Displayed by the Human Hematopoietic KG1a Progenitor Cells Line and by Primary Human CD34<sup>+</sup> Cells," *Blood*, **92**:3616-3623
131. C.H. Schilling, S. Schuster, B.O. Palsson, and R. Heinrich (1999), "Metabolic Pathway Analysis: Basic Concepts and Scientific Applications in the Post-Genomic Era," *Biotechnology Progress*, **15**: 296-303.
132. J.S. Edwards and B.O. Palsson (1999), "Systems Properties of the *Haemophilus influenzae* Rd Metabolic Genotype," *The Journal of Biological Chemistry*, **274**: 17410-17416.
133. C.H. Schilling, J.S. Edwards and B.O. Palsson (1999), "Toward Metabolic Phenomics: Analysis of Genomic Data Using Flux Balances," *Biotechnology Progress* **15**: 288-295.
134. D.J. Oh, Gyun Min Lee, Karl Francis, and B.O. Palsson (1999), "Phototoxicity of the Fluorescent Membrane Dyes PKH2 and PKH26 on the Human Hematopoietic KG1a Progenitor Cell Line," *Cytometry*, **36**: 312-318.

135. S. Huang, P. Law, K. Francis, B.O. Palsson, and A.D. Ho (1999), "Symmetry of Initial Cell Divisions Among Primitive Hematopoietic Progenitors is Independent of Ontogenic Age and Regulatory Molecules," *Blood* **94**: 2595-2604.
136. W. Holloway, A. Martinez, D.J. Oh, K. Francis, R. Ramakrishna, and B.O. Palsson (1999), "Key adhesion molecules are present on long podia extended by hematopoietic cells," *Cytometry* **37**: 171-177.
137. Francis, K., Oh, D.J., Martinez, A., Palsson, B.Ø., "Response to Tenopodia Formation by Blasts in Bone Marrow Aspirate Smears from a Patient with Acute Monoblastic Leukemia", *Blood*, **94**: 3955, (1999).
138. C.H. Schilling, D. Letscher, and B.O. Palsson, "Theory for the Systemic Definition of Metabolic Pathways and their use in Interpreting Metabolic Function from a Pathway-Oriented Perspective," *J. Theoret. Biol.*, **203**: 229-248 (2000).
139. C.H. Schilling and B.O. Palsson, "Assessment of the Metabolic Capabilities of *Haemophilus influenzae* Rd through a Genome-Scale Pathway Analysis," *J. Theoret. Biol.*, **203**: 249-283 (2000).
140. D.J. Oh, A.R. Martinez, G.M. Lee, K. Francis, and B.O. Palsson (2000). "Extension of Osmolality-Induced Podia Is Observed from Fluorescently Labeled Hematopoietic Cell Lines in Hyperosmotic Medium," *Cytometry*, **40**: 109-118.
141. D.J. Oh, A.R. Martinez, G.M. Lee, K. Francis, and B.O. Palsson (2000). "Inter-cellular Adhesion can be Visualized Using Fluorescently Labeled Fibrosarcoma HT1080 Cells Cocultured with Hematopoietic Cell Lines or CD34<sup>+</sup> Enriched Human Mobilized Peripheral Cells," *Cytometry*, **40**: 119-125.
142. J.S. Edwards and B.O. Palsson, "The *Escherichia coli* MG1655 *in silico* metabolic genotype; Its definition, characteristics, and capabilities," *Proc. Natl. Acad. Sci. (USA)*, **97**: 5528-5523 (2000).
143. J.S. Edwards and B.O. Palsson, "Metabolic flux balance analysis and the *in silico* analysis of *Escherichia coli* K-12 gene deletions," *BMC Bioinformatics*, **1**: 1 (2000).
144. J.S. Edwards and B.O. Palsson, "Robustness Analysis of the *Escherichia coli* Metabolic Network," *Biotechnology Progress*, **16**: 927-939 (2000).
145. R. Ramakrishna, J.S. Edwards, A. McCulloch, B.O. Palsson, "Flux Balance Analysis of Mitochondrial Energy Metabolism: Consequences of Systemic Stoichiometric Constraints," *The American Journal of Physiology: Regulatory, Integrative and Comparative Physiology*, **280**: R695-R704 (2001).
146. N. Jamshidi, J.S. Edwards, T. Fahland, G.M Church, and B.O. Palsson, "Dynamic simulation of the human red blood cell metabolic network," *Bioinformatics*, **17**: 286-287, (2001).
147. M.W. Covert, C.H. Schilling, I. Famili, J.S. Edwards, I.I. Goryanin, E. Selkov, and B.O. Palsson, "Metabolic modeling of microbial stains *in silico*," *Trends in Biochemical Sciences*, **26**: 179-186 (2001).
148. J.S. Edwards, R.U. Ibarra, and B.O. Palsson, "In *silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data," *Nature Biotechnology*, **19**: 125-130 (2001).
149. C.H. Schilling, J. S. Edwards, D. Letscher, and B.O. Palsson, "Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems," *Biotechnology and Bioengineering*, **71**: 286-306 (2001).
150. M.W. Covert, C.H. Schilling, and B.O. Palsson, "Regulation of gene expression in flux balance models of metabolism," *Journal of Theoretical Biology*, **213**: 73-88 (2001).
151. K. Francis, B.O. Palsson, J. Donahue, S. Fong, and E. Carrier, "Murine Sca-1<sup>+</sup>Lin<sup>-</sup> cells and human KG1a cells exhibit multiple pseudopod morphologies during migration," *Experimental Hematology*, **30**: 460-463, (2002).
152. J.S. Edwards, R. Ramakrishna, and B.O. Palsson, "Characterizing the Metabolic Phenotype: A Phenotype Phase Plane Analysis," *Biotechnology and Bioengineering*, **77**: 27-36, (2002).
153. J.A. Papin, N.D. Price, J.S. Edwards, and B.O. Palsson, "The Genome-Scale Metabolic Extreme Pathway Structure in *Haemophilus influenzae* Shows Significant Network Redundancy," *Journal of Theoretical Biology*, **215**: 67-82 (2002).
154. N.D. Price, J.A. Papin, and B.O. Palsson, "Determination of Redundancy and Systems Properties of the Metabolic Network of *Helicobacter pylori* Using Genome-Scale Extreme Pathway Analysis," *Genome Research*, **12**: 760-769 (2002).

155. J.S. Edwards, M. Covert, and B.O. Palsson, "Metabolic Modeling of Microbes: the Flux Balance Approach," *Environmental Microbiology*, **4**: 133-140, (2002).
156. M.W. Covert and B.O. Palsson, "Transcriptional Regulation in Constraints-Based Metabolic Models of *Escherichia coli*," *Journal of Biological Chemistry*, **277**: 28058-28064 (2002).
157. S.J. Smith and B.O. Palsson, "Extreme Pathway Analysis of Human Red Blood Cell Metabolism," *Biophysical Journal*, **83**: 808-818 (2002).
158. C.H. Schilling, M.W. Covert, I. Famili, G.M. Church, J.S. Edwards, and B.O. Palsson, "Genome-Scale Metabolic Model of *Helicobacter pylori* 26695," *Journal of Bacteriology*, **184**: 4582-4593 (2002).
159. T.E. Allen and B.O. Palsson, "Sequence-based analysis of metabolic demands for protein synthesis in prokaryotes," *Journal of Theoretical Biology*, **220**: 1-18 (2003).
160. M.W. Covert and B.O. Palsson, "Constraints-Based Models: Regulation of Gene Expression Reduces the Steady-State Solution Space", *Journal of Theoretical Biology*, **221**:309-325 (2003).
161. N.D. Price, I. Famili, D.A. Beard, and B.O. Palsson, "Extreme Pathways and Kirchhoff's Second Law", *Biophysical Journal*, **83**: 2879-2882 (2002).
162. N.D. Price, J.L. Reed, J.A. Papin, I. Famili, and B.O. Palsson, "Analysis of metabolic capabilities using singular value decomposition of extreme pathway matrices," *Biophysical Journal*, **84**:794-804 (2003).
163. J.A. Papin, N.D. Price and B.O. Palsson, "Extreme Pathway Lengths and Reaction Participation in Genome-Scale Metabolic Networks," *Genome Research*, **12**:1889-1900 (2002).
164. K.J. Kauffman, J.D. Pajewski, N. Jamshidi, B.O. Palsson, and J.S. Edwards, "Description and analysis of metabolic connectivity and dynamics in the human red blood cell," *Biophysical Journal*, **83**:646-662 (2002).
165. N. Jamshidi, S.J. Wiback and B.O. Palsson, "In silico model-driven assessment of the effects of single nucleotide Polymorphisms (SNPs) on human red blood cell metabolism," *Genome Research*, **12**:1687-1692 (2002).
166. R.U. Ibarra, J.S. Edwards, and B.O. Palsson, "*Escherichia coli* K-12 undergoes adaptive evolution to achieve *in silico* predicted optimal growth," *Nature*, **420**: 186-189 (2002).
167. Lee, G.M., Fong, S., Francis, K., Palsson, B.Ø., "Characterization and Efficacy of PKH26 as a Probe to Study the Replication History of the Human Hematopoietic KG1a Progenitor Cell Line", *In Vitro Cellular and Developmental Biology: Animal*, **38**(2): 90-96 (2002).
168. Francis, K., Lee, G.M., Palsson, B.Ø., "Characterization of the KG1a Cell Line for use in a Cell Migration Based Screening Assay", *Biotechnology and Bioengineering*, **7**(3): 178-184 (2002).
169. J. Forster, I. Famili, P.C. Fu, B.O. Palsson and J. Nielsen, "Genome-Scale Reconstruction of the *Saccharomyces cerevisiae* Metabolic Network," *Genome Research*, **13**:244-253 (2003).
170. N.D. Price, J.A. Papin, C.H. Schilling and B.O. Palsson, "Genome-scale Microbial In Silico models: The Constraints-based Approach," *Trends in Biotechnology*, **21**: 162-169 (2003).
171. A. Raghunathan and B.O. Palsson, "Scalable method to determine mutations that occur during adaptive evolution of *Escherichia coli*," *Biotechnology Letters*, **25**:435-441 (2003).
172. Papin, J.A., Price, N.D., Wiback, S.J., Fell, D.A., and Palsson, B.O., "Metabolic Pathways in the Post-Genome Era," *Trends in Biochemical Sciences*, **28**:250-258 (2003).
173. Reed, J.L. and Palsson, B.O., "Thirteen Years of Building Constraints-Based *in silico* Models of *Escherichia coli*," *Journal of Bacteriology*, **185**:2692-2699 (2003).
174. Famili, I. and Palsson, B.O., "The Convex Basis of the Left Null Space of the Stoichiometric Matrix Leads to the Definition of Metabolically Meaningful Pools," *Biophysical Journal*, **85**:16-26 (2003).
175. Wiback, S.J., Mahadevan, R., and Palsson, B.O., "Reconstructing metabolic flux vectors from extreme pathways: defining the  $\alpha$ -Spectrum," *Journal of Theoretical Biology*, **24**:313-324 (2003).
176. Forster, J., Famili, I., Palsson, B.O. and Nielsen, J., "Large-scale Evaluation of *in silico* Gene Deletions in *Saccharomyces cerevisiae*," *OMICS*, **7**: 193-202 (2003).
177. Famili, I. and Palsson, B.O., "Systemic metabolic reactions are obtained by singular value decomposition of genome-scale stoichiometric matrices," *Journal of Theoretical Biology*, **224**: 87-96 (2003).

178. Price, N.D., Reed, J.L., Papin, J.A., Wiback, S.J., and Palsson, B.O., "Network-based Analysis of Regulation in the Human Red Blood Cell," *Journal of Theoretical Biology*, **225**: 1985-1994 (2003).
179. Herrgard, M., Covert, M.W., and Palsson, B.O., "Reconciling Gene Expression Data with Known Genome-scale Regulatory Network Structure," *Genome Research*, **13**: 2423-2434, (2003).
180. Reed, J.L., Vo, T.D., Schilling, C.H., and Palsson, B.O., "An expanded genome-scale model of *Escherichia coli* K-12 (JUR904 GSM/GPR)," *Genome Biology*, **4**:R54.1-R54.12 (2003).
181. Allen, T.E., Herrgard, M.J., Mingzhu, L., Yu, Q., Glasner, J.D., Blattner, F.R., and Palsson, B.O., "Genome-scale analysis of the uses of the *Escherichia coli* genome: a model-driven analysis of heterogeneous datasets," *Journal of Bacteriology*, **185**: 6392-6399, (2003).
182. Ibarra, R.U., Fu, P., Palsson, B.O., DiTonno, J.R., Edwards, J.S., "Quantitative Analysis of *E. coli* Metabolic Phenotypes within the Context of Phenotypic Phase Planes." *J. Mol. Microbiol. Biotechnol.*, **6**:101-108, (2003)
183. Fong, S.S. Marciniak, J.Y., and Palsson, B.O., "Description and Interpretation of Adaptive Evolution of *Escherichia coli* K12 MG1655 Using a Genome-scale *in silico* Model," *Journal of Bacteriology*, **185**: 6400-6408, (2003).
184. Kolker, E., Purvine, S., Galperin, M.Y., Stolyar, S., Goolett, D.R., Nesvizhskii, A.I., Keller, A., Xie, T., Eng, J.K., Yi, E., Hood, L., Picone, A.F., Chernyl, T., Tjaden, B.C., Siegel, A.F., Reilly, T.J., Makarova, K.S., Palsson, B.O., and Smith, A.L., "Initial Proteome Analysis of Model Microorganism *Haemophilus Influenzae* Strain Rd KW20," *Journal of Bacteriology*, **185**:4593-4602 (2003).
185. Famili, I., Forster, J., Nielsen, J., and Palsson, B.O., "*Saccharomyces cerevisiae* Phenotypes can be Predicted Using Constraint-based Analysis of a Genome-Scale Reconstructed Metabolic Network," *PNAS*, **100**: 13134-13139, (2003).
186. Covert, M.W. Famili, I., and Palsson, B.O., "Identifying the Constraints that Govern Cell Behavior: A Key to Converting Conceptual to Computational Models in Biology?," *Biotechnology and Bioengineering*, **84**: 763-772, (2003).
187. Palsson, B.Ø., Price, N.D., Papin, J.A., "Development of network-based pathway definitions: the need to analyze real metabolic networks", *Trends in Biotechnology*, **21**(5): pp. 195-198 (2003).
188. Papin, J.A. and Palsson, B.O., "Topological Analysis of Mass-Balanced Signaling Networks: A Framework to Obtain Emergent Properties Including Crosstalk," *Journal of Theoretical Biology*, **227**, 283-297(2004).
189. Palsson, B.O., "In silico biotechnology: Era of reconstruction and interrogation," *Current Opinion in Biotechnology*, **15**: 50-51 (2004).
190. Herrgard, M.J., "Reconstruction of microbial transcription regulatory networks," *Current Opinion in Biotechnology*, **15**: 70-77 (2004).
191. Wiback, S.J., Mahadevan, R., and Palsson, B.Ø., "Using Metabolic Flux Data to Further Constrain the Metabolic Solution space and Predict Internal Flux Patterns: The *Escherichia coli* a-spectrum", *Biotechnology & Bioengineering*, **86**: 317-331, (2004).
192. Bell, S.L. and Palsson, B.Ø., "Phenotype Phase Plane Analysis Using Interior Point Methods", *Computers & Chemical Engineering*, **29**: 481-486, (2005).
193. Wiback, S.J., Famili, I., Greenberg, H.J., and Palsson, B.Ø., "Monte Carlo Sampling Can Be Used to Determine the Size and Shape of the Steady State Flux Space", *Journal of Theoretical Biology*, **228**: 437-444, (2004).
194. Covert, M.W., Knight, E.M., Reed, J.L., Herrgard, M.J., and Palsson, B.Ø., "Integrating high-throughput and computational data elucidates bacterial networks", *Nature*, **429**: 92-96, (2004).
195. Papin, J.A. and Palsson, B.Ø., "The JAK-STAT Signaling Network in the Human B-Cell: An Extreme Signaling Pathway Analysis", *Biophysical Journal*, **87**: 37-46, (2004).
196. Duarte, N.C., Herrgard, M.J. and Palsson, B.Ø., "Reconstruction and Validation of *Saccharomyces cerevisiae* iND750, a Fully Compartmentalized Genome-scale Metabolic Model", *Genome Research*, **14**: **1298-1309**, (2004).
197. Papin, J.A., Stelling, J., Price, N.D., Klamt, S., Schuster, S., and Palsson, B.Ø., "Comparison of Network-Based Pathway Analysis Methods," *Trends in Biotechnology*, **22**: 400-405, (2004).



198. Price, N.P., Schellenberger, J., and Palsson, B.O., "Uniform Sampling of Steady State Flux Spaces: Means to Design Experiments and to Interpret Enzymopathies," *Biophysical Journal*, **87**: 2172-2186, (2004).
199. Fong, S.S. and Palsson, B.Ø., "Metabolic gene-deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes," *Nature Genetics*, **36**(10): 1056-58, (2004).
200. Duarte, N.C., Palsson, B.Ø., and Fu, P., "Integrated Analysis of Metabolic Phenotypes in *Saccharomyces cerevisiae*," *BMC Genomics*, **5**:63 (2004).
201. Papin, J.A., Reed, J.L. and Palsson, B.Ø., "Hierarchical thinking in network biology: the unbiased modularization of biochemical networks," *Trends in Biochemical Sciences*, **29**: 641-647, (2004).
202. Honisch, C., Raghunathan, A., Cantor, C.R., Palsson, B.Ø. and van den Boom, D., "High-throughput Mutation Detection Underlying Adaptive Evolution of *Escherichia coli*-K12," *Genome Research*, **14**(12): 2495-2502 (2004).
203. Koller, M.R., Hanania, E.G., Stevens, J., Eisfeld, T.M., Sasaki, G.C., Fleck, A., and Palsson, B.O., "High-throughput laser-mediated *in situ* cell purification with high purity and yield," *Cytometry*, **61A**: 153-161 (2004).
204. Mahadevan, R. and Palsson, B.Ø., "Properties of Metabolic Networks: Structure vs. Function," *Biophysical Journal*, **88**: L7-L9 (2005).
205. Thiele, I., Price, N.D., Vo, T.D. and Palsson, B.Ø., "Candidate metabolic network states in human mitochondria: Impact of diabetes, ischemia, and diet" *Journal of Biological Chemistry*, **280**: 11683-11695 (2005).
206. Papin, J., Hunter, T., Palsson, B.Ø., and Subramanian, S., "Reconstruction of cellular signaling networks and analysis of their properties," *Nature Reviews Molecular Cell Biology*, **6**(2):99-111 (2005).
207. Bell, S.L. and Palsson, B.Ø., "ExPA: A Program for Calculating Extreme Pathways in Biochemical Reaction Networks," *Bioinformatics*, **21**: 481-486, (2005).
208. Famili, I., Mahadevan, R. and Palsson, B.Ø., "K-Cone Analysis: Determining All Candidate Values for Kinetic Parameters on a Network-scale," *Biophysical Journal*, (2005).
209. Raghunathan, A., Price, N.D., Galperin, M.Y., Makarova, K.S., Purvine, S., Picone, A.F., Cherny, T., Xie, T., Reilly, T.J., Munson, Jr., R., Tyler, R.E., Akerley, B.J., Smith, A.L., Palsson, B.Ø., and Kolker, E. "In Silico Metabolic Model and Protein Expression of *Haemophilus influenzae* strain Rd KW20 in Rich Medium", *OMICS*, **8**(1): 25-42 (2004).
210. Price, N.D., Reed, J.L. and Palsson, B.Ø., "Genome-scale Models of Microbial Cells: Evaluating the consequences of constraints", *Nature Reviews Microbiology*, **2**:886-897 (2004).
211. Vo, T.D., Greenberg, H.J., and Palsson, B.O., "Reconstruction and functional characterization of the human mitochondrial metabolic network based on proteomic and biochemical data", *Journal of Biological Chemistry*, **279**(38): 39532-40 (2004).
212. Reed, J.L. and Palsson, B.Ø., "Genome-scale *in silico* models of *E. coli* have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states", *Genome Research*, **14**:1797-1805(2004).
213. Becker, S.A. and Palsson, B.Ø., "Genome-scale reconstruction of the metabolic network in *Staphylococcus aureus* N315: an initial draft to the two-dimensional annotation", *BMC Microbiology*, **5**: 8 (2005).
214. Fong, S.S., Burgard, A.P., Herring, C.D., Knight, E.M., Blattner, F.R., Maranas, C.D., and Palsson, B.Ø., "In Silico Design and Adaptive Evolution of *Escherichia coli* for Production of Lactic Acid", *Biotechnology and Bioengineering*, **91**: 643-648 (2005).
215. Thiele, I., Vo, T.D., Price, N.D. and Palsson, B.Ø., "An Expanded Metabolic Reconstruction of *Helicobacter pylori* (J1341 GSM/GPR): An *in silico* genome-scale characterization of single and double deletion mutants", *Journal of Bacteriology*, **187**: 5818-5830 (2005).
216. Herring, C.D., Raffaele, M., Allen, T.E., Kanin, E., Landick, R., Ansari, A.Z., and Palsson, B.Ø., "Immobilization of *Escherichia coli* RNA polymerase and location of binding sites using chromatin immunoprecipitation and microarrays", *Journal of Bacteriology*, **187**: 6166-6174 (2005).
217. Fong, S.S., Joyce, A.J., and Palsson, B.Ø., "Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different expression states", *Genome Research*, **15**(10):1365-1372 (2005).

218. Barrett, C.B., Herring, C.D., Reed, J.L., and Palsson, B.Ø., "The Global Transcriptional Regulatory Network for Metabolism in *Escherichia coli* Exhibits Few Dominant Functional States", *PNAS*, **102**(52):19103-8 (2005).
219. Hanania, E.G., Fleck, A., Stevens, J., Bodzin, L., and Palsson, B.O., "Automated *in situ* measurement of cell-specific antibody secretion and laser-mediated purification for rapid cloning of highly-secreting producers," *Biotechnology and Bioengineering*, **91**: 872-876 (2005).
220. Cho, B.K., Knight, E.M., and Palsson, B.Ø., "PCR-based tandem epitope tagging system for *E. coli* genome engineering", *Biotechniques*, **40**:67-70 (2006).
221. Reed, J.L., Famili, I., Thiele, I., and Palsson, B.Ø., "Towards multidimensional genome annotation", *Nature Reviews Genetics*, **7**(2):130-41 (2006).
222. Feist, A.M., Scholten, J.C.M., Palsson, B.Ø., Brockman, F.J., and Ideker, T., "Modeling methanogenesis with a genome-scale metabolic reconstruction of *Methanosarcina barkeri*", *Molecular Systems Biology*, **2**(1): msb4100046-E1-E14 (2006).
223. Joyce A.R. and Palsson B.Ø. "The model organism as a system: integrating 'omics' data sets", *Nature Reviews Molecular Cell Biology*, **7**: 198-210 (2006)
224. Herrgard, M.J., Lee, B.-S., Portnoy, V., Palsson, B.Ø., "Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in *Saccharomyces cerevisiae*", *Genome Research*, **16**: 627-635 (2006).
225. Barrett, C.L. and Palsson, B.Ø., "Iterative reconstruction of transcriptional regulatory networks: an algorithmic approach", *PLoS Computational Biology*, **2**: 5 (2006).
226. Becker, S.A., Price, N.D., Palsson, B.Ø., "Metabolite coupling in genome-scale metabolic networks", *BMC Bioinformatics*, **7**:111 (2006).
227. Barrett, C.L., Price, N.D. Palsson, B.Ø., "Network-level analysis of metabolic regulation in the human red blood cell using random sampling and singular value decomposition", *BMC Bioinformatics* **7**:132 (2006)
228. Cho, B.K., Knight, E.M., Palsson, B.Ø., "Transcriptional regulation of the fad regulon genes of *Escherichia coli* by ArcA", *Microbiology*, **152**: 2207-2219 (2006).
229. Jamshidi N, Palsson BO. "Systems biology of the human red blood cell", *Blood Cells Mol Dis*, **36**(2):239-47 (2006).
230. Joyce, A. R., Reed, J.L., White, A., Edwards, R., Osterman, A., Baba, T., Lesley, S.A., Palsson, B.O., Agarwalla, S., "Experimental and computational assessment of conditionally essential genes in *E. coli*", *Journal of Bacteriology*, **188**: 8259-8271 (2006).
231. Gianchandani E. P., Papin J. A., Price N. D., Joyce A. R., Palsson B. O., "Matrix formalism to describe functional States of transcriptional regulatory systems," *PLoS Comput Biol*, **11**;2(8):e101 (2006).
232. Vo, T. D., Lim, S. K., Lee, W. N. P., Palsson, B. O., "Isotopomer Analysis of Cellular Metabolism in Tissue Culture: A Comparative study between the pathway and network-based methods", *Metabolomics*, **2**: 243-256 (2006).
233. Jamshidi, N., Palsson, B.Ø., "Systems biology of SNPs," *Molecular Systems Biology*, **2**: doi:10.1038/msb4100077 (2006).
234. Herrgard, M.J., Fong, S.F., Palsson, B.Ø., "Systematic identification of genome-scale metabolic models using experimentally measured flux profiles.", *PLoS Computational Biology*, **2**: e72 (2006).
235. Hua, Q., Joyce, A.R., Fong, S.S., Palsson, B.Ø., "Metabolic Analysis of Adaptive Evolution for In Silico-Designed Lactate-Producing Strains," *Biotechnology and Bioengineering*, **95**: 992-1002 (2006).
236. Reed, J. R., Patel, T. R., Chen, K. H., Joyce, A. R., Applebee, M. K., Herring, C. D., Bui, O. T., Knight, E. K., Fong, S. S., Palsson, B. O., "Systems Approach to Refining Genome Annotation", *PNAS*, **03**(46):17480-4 (2006).
237. Herring, C. D., Raghunathan, A., Honisch, C., Patel, T., Applebee, M. K., Joyce, A. J., Albert, T. J., Blattner, F. R., van den Boom, D., Cantor, C. R., Palsson, B. O., "Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale", *Nature Genetics*, **38**: 1046-1412; (2006)
238. Price, N.D., Thiele, I., Palsson B.Ø. "Candidate States of *Helicobacter pylori*'s Genome-Scale Metabolic Network upon Application of "Loop Law" Thermodynamic Constraints" *Biophys J*, **90**:3919-28 (2006)

239. Clark, I.B., Hanania, E.G., Stevens, J., Gallina, M., Fleck, A., Brandes, R., Palsson, B.O., and Koller, M.R., "Optoinjection for efficient targeted delivery of a broad range of compounds and macromolecules into diverse cell types," *Journal of Biomedical Optics*, **11**: 014034-1-014034-8 (2006).
240. Fong, S.S., Nanchen, A., Palsson, B.O., and Sauer, U., "Latent Pathway Activation and Increased Pathway Capacity Enable *Escherichia coli* Adaptation to Loss of Key Metabolic Enzymes," *Journal of Biological Chemistry*, **281**:8024-8033 (2006).
241. Mahadevan, R., Bond, D.R., Butler, J.E., Esteve-Nunez, A., Coppi, M.V., Palsson, B.O., Schilling, C.H., and Lovley, D.R., "Characterization of metabolism in the Fe(II)-reducing organism *Geobacter sulfurreducens* by constraint-based modeling," *Appl. Environ. Microbiol.*, **72**: 1558-1568 (2006).
242. Allen, T.E., Price, N.D., Joyce, A.R., and Palsson, B.Ø., "Long-range patterns in prokaryotic genome sequences indicate significant chromosomal organization," *PLoS Computational Biology*, **13**:2(1):e2 (2006).
243. Duarte, N.D., Becker, S.A., Jamshidi, N., Thiele, I., Mo, M. L., Vo, T. D., Srivas, R., Palsson, B. O., "Global reconstruction of the human metabolic network based on genomic and bibliomic data", *Proc. Nat'l Acad. Sci.* **104**:1777-1782 (2007).
244. Becker, S.A., Feist, A.M., Mo, M. L., Hannum, G., Palsson, B.O., Herrgard, M.J. "Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox" *Nat. Protocols* **2**, 727-738 (2007).
245. Vo, T.D., Lee, W. N. P., Palsson, B. O. "Systems analysis of energy metabolism elucidates the 3 affected respiratory chain complex in Leigh's syndrome", *Mol Genet Metab.* doi:10.1016/j.ymgme.2007.01.012 (2007).
246. Vo, T.D., Palsson, B.O. "Building the power house: recent advances in mitochondrial studies through proteomics and systems biology" *Am J Physiol Cell Physiol.* 292(1):C164-77 (2007)
247. Feist, A.M., Henry, C.S., Reed, J.L., Krummenacker, M., Joyce, A.R., Karp, P.D., Broadbelt, L.J., Hatzimanikatis, V. & Palsson, B.O. "A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information" *Molecular Systems Biology* **3**: 121 (2007).
248. Mo, M.L., Jamshidi, N., Palsson, B.Ø., "A Genome-scale, Constraint-based Approach to Systems Biology of Human Metabolism," *Molecular Biosystems* **3**: 598.
249. Oh, Y-K., Palsson, B.Ø., Park, S.M., Schilling, C.H., Mahadevan, R. "Genome-scale Reconstruction of Metabolic Network in *Bacillus subtilis* Based on High-throughput Phenotyping and Gene Essentiality Data" *J Biol Chem.* 10.1074/jbc.M703759200 (2007).
250. Jamshidi, N, Palsson, B.Ø. "Investigating the Metabolic Capabilities of *Mycobacterium tuberculosis* H37Rv Using the in silico Strain INJ661 and Proposing Alternative Drug Targets" *BMC Systems Biology* **1**:26 (2007).
251. Herring, C.S., Palsson, B.Ø "An evaluation of Comparative Genome Sequencing (CGS) by comparing two previously-sequenced bacterial genomes" *BMC Genomics* **8**:274 (2007).
252. Resendis-Antonio O., Reed J.L., Encarnacion S., Collado-Vides J. and Palsson B.Ø. "Metabolic Reconstruction and Modeling of Nitrogen Fixation in *Rhizobium etli*" *PLoS Comp. Biol.* **3**: 10 (2007).
253. Shlomi, T., Herrgard, M., Portnoy, V., Naim, E., Palsson, B.Ø., Sharan, R., Ruppin, E., Systematic condition-dependent annotation of metabolic genes., *Genome Res.*, **17**(11): 1626-33 (2007).
254. Cho, B.K., Charusanti, P., Herrgard, M.J. and Palsson, B.Ø., "Microbial Regulatory and Metabolic Networks" *Current Opinion in Biotechnology* **18**: 4 (2007).
255. Yeung, M., Thiele, I., and Palsson, B.Ø., Estimation of the number of extreme pathways for metabolic networks," *BMC Bioinformatics*, **8**(1): 363 (2007).
256. Hua, Q., Joyce, A.R., Palsson, B.Ø., Fong, S.S., Metabolic characterization of *Escherichia coli* strains adapted to growth on lactate., *Appl Environ Microbiol.*, **73**(14): 4639-47 (2007).
257. Anderson, M.J., DeLaBarre, B., Raghunathan, A., Palsson, B.O., Brunger, A.T., Quake, S.R., "Crystal Structure of a Hyperactive *Escherichia coli* Glycerol Kinase Mutant Gly230→ Asp Obtained Using Microfluidic Crystallization Devices," *Biochemistry* **46**: 5722-5731 (2007).
258. Cho, B.K., Knight, E.M., Barrett, C.L. and Palsson, B.Ø., "Genome-wide analysis of Fis binding in *Escherichia coli* indicates a causative role for A-tracts", Submitted.

259. Becker, S.A., Palsson, B.O., Three factors underlying incorrect *in silico* predictions of essential metabolic genes. *BMC Syst Biol.*, **2**(1): 14 (2008).
260. Jamshidi, N. and Palsson, B.O., "Formulating genome-scale kinetic models in the post-genome era," *Mol. Sys. Biol.*, **4**: 171-180 (2008).
261. Cho, B.K., Knight, E.M., and Palsson, B.Ø., Genome-wide Identification of Protein Binding Locations using Chromatin Immunoprecipitation coupled with Microarrays., *Methods Mol Biol.*, **439**: 131 (2008).
262. Joyce, A.R. and Palsson, B.Ø., Predicting gene essentiality using genome-scale *in silico* models. *Methods Mol Biol*, **416**: 433-57(2008).
263. Feist, A.M. and Palsson, B.Ø., The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*. *Nature Biotechnology*, **26**(6): 659 - 667 (2008).
264. Becker, S.A. and Palsson, B.Ø., Context-specific metabolic networks are consistent with experiments., *PLoS Comput Biol*, **4**(5): e1000082 (2008).
265. Applebee M.K., Herrgård M.J. and Palsson, B.Ø., Impact of Individual Mutations on Increased Fitness in Adaptively-evolved strains of *Escherichia coli*., *J Bacteriol*, EPub (2008).
266. Shlomi, T., Cabili, M.N., Herrgård, M.J., Palsson, B.O. and Ruppin, E., Network-based prediction of human tissue-specific metabolism, *Nature Biotechnology*, doi:10.1038/nbt.1487 (2008).
267. Lee, J., Yun, H., Feist, A.M., Palsson, B.O., and Lee, S.Y., "Genome-scale reconstruction and *in silico* analysis of the *Clostridium acetobutylicum* ATCC 824 metabolic network", *Appl. Microbiol. Biotechnol.* **80**: 849-862 (2008).
268. Feist, A.M., Herrgård, M.J., Thiele, I., Reed, J.L., Palsson, B.Ø., "Reconstruction of Biochemical Networks in Microbial Organisms" *Nature Reviews Microbiology*, **7**: 129-143(2009).
269. Jamshidi, N. and Palsson, B.O., "Top-Down Analysis of Temporal Hierarchy in Biochemical Reaction Networks," *PLoS Computational Biology*, **4**(9): e1000177. doi:10.1371/journal.pcbi.1000177 (2008).
270. Nogales, J., Palsson, B.Ø., and Thiele, I., A genome-scale metabolic reconstruction of *Pseudomonas putida* KT2440: iJN746 as a cell factory., *BMC Syst Biol*, **2**:79.doi:10.1186/1752-0509-2-79 (2008).
271. Herrgård, M. J., Swainston, N., Dobson, P., Dunn, W. B., Arga, K. Y., Arvas, M., Blüthgen, N., Borger, S., Costenoble, R., Heinemann, M., Hucka, M., Le Novère, N., Li, P., Liebermeister, W., Mo, M. L., Oliveira, A. P., Petranovic, D., Pettifer, S., Simeonidis, E., Smallbone, K., Spasić, I., Weichert, D., Brent, R., Broomhead, D. S., Westerhoff, H. V., Kirdar, B., Penttilä, M., Klipp, E., Palsson, B. Ø., Sauer, U., Oliver, S. G., Mendes, P., Nielsen, J. & Kell, D. B., A consensus yeast metabolic network obtained from a community approach to systems biology., "Nature Biotechnology", **26**: 1155-1160 (2008).
272. Mo, M.L. and Palsson, B.Ø. "Understanding human metabolic physiology: A genome-to-systems approach," *Trends in Biotechnology*. doi:10.1016/j.tibtech.2008.09.007 (2008).
273. Schellenberger, J. and Palsson, B.Ø. "The use of randomized sampling for analysis of metabolic networks", *JBC Minireview*. (Accepted).
274. Cho, B.K., Barrett, C.L., Knight, E.M., Park, Y.S., and Palsson B.Ø. Genome-scale reconstruction of Lrp-regulatory network in *Escherichia coli*, *Proc Natl Acad Sci USA*, Published Online (2008).
275. Portnoy, V.A., Herrgård, M.J., Palsson, B.Ø. "Aerobic Fermentation of D-Glucose by an Evolved Cytochrome Oxidase Deficient *Escherichia coli* Strain," *Appl Environ Microbiol.* **74**: 7561-7569 (2008).
276. Sigurdsson, M., Jamshidi, N., Jonsson, J.J. and Palsson, B.O., "Genome-scale network analysis of imprinted human metabolic genes," *Epigenetics*, **4**: (2009).
277. Li, F., Thiele, I., Jamshidi, N. and Palsson, B.O., "Identification of Potential Pathway Mediation Targets in Toll-like Receptor Signaling," *PLOS Comp Biol*. **5**: e1000292. doi:10.1371/journal.pcbi.1000292 (2009).
278. Thiele, I., Jamshidi, N., Fleming, M.T., and Palsson, B.Ø., "Genome-scale reconstruction of *E. coli* 's transcriptional and translational machinery: A knowledge-base, its mathematical formulation, and its functional characterization," *PLoS Comput Biol*, **5**: e1000312 (2009).

279. Barrett, C.L., Herrgard, M.J., and Palsson, B.Ø., "Decomposing complex reaction networks using random sampling, principal component analysis, and basis rotation," *BMC Systems Biology*, **3**:30 (2009).
280. Lewis, N.E., Cho, B.K., Knight, E.M., and Palsson, B.O., Gene Expression Profiling and the Use of Genome-scale *In Silico* Models of *Escherichia coli* for Analysis: Providing Context for Content, *J Bacteriol*, **Accepted** (2009).
281. Raghunathan, A., Reed, J., Shin, S., Palsson, B.O., Daefler, S., Constraint-based analysis of metabolic capacity of *Salmonella typhimurium* during hostpathogen interaction. *BMC Systems Biology*, **3**:38 (2009).

## PUBLICATIONS -- Non-Refereed

1. J. B. Dressman, S. S. Ozturk, and B. O. Palsson (1988), "Predicting Release Kinetics from Enteric Coated Dosage Forms," *Pharmaceutical Technology*, September 1988
2. B. O. Palsson, A. Narang, and A. Joshi (1988), "Computer Model of Human Erythrocyte Metabolism," in Seventh International Conference on Red Cell Metabolism and Function, G. Brewer, Editor
3. M.F. Clarke, F.Y. El-Awar, R.M. Schwartz, B.O. Palsson, and S.G. Emerson (1993), "Retroviral Gene Transfer into Human Hematopoietic Cells Using Rapidly Perfused Long-Term Bone Marrow Cultures," *The Cancer Bulletin*, **45**:2, 153-158.
4. M.R. Kolter, R.J. Maher and B. O. Palsson (1996), "Hydrocortisone and SCF interact to Influence the Regulation of the Hematopoietic Cell Development by the Bone Marrow Microenvironment," Keystone Bone Marrow Meeting Proceedings
5. G.M. Lee, S. Fong, K. Francis, D.J. Oh, and B.O. Palsson (2000), Letter to the Editor: "In situ labeling of adherent cells with PKH26," *In Vitro Cell. Dev. Biol.-Animal* **36**: 4-6.
6. J.S. Edwards and B.O. Palsson (2000), "Letter to the Editor: Multiple Steady States in Kinetic Models of Red Cell Metabolism," *J. theor. Biol.* **207**: 125-127.
7. K. Francis and B.O. Palsson (2002), Comment on "Time Dependent Concentration Profile of Secreted Molecules in the Intercellular Signaling," *Journal of the Physical Society of Japan*, **71**: 1005-1006.

## COMMENTARIES

1. B.O. Palsson, "What Lies Beyond Bioinformatics," *Nature Biotechnology*, **15**:3-4 (1997).
2. B.O. Palsson, S. Thorgeirsson, "Decoding Developments in Iceland," *Nature Biotechnology*, **17**:5, p.407 (1999).
3. B.O. Palsson (2000), "The challenges of *in silico* biology," *Nature Biotechnology*, **18**: 1147-1150.
4. M.W. Covert, B.O. Palsson and C.H. Schilling, "A more-palatable H-pylori: Iterative model-building through the peer-review process," *ASM News*, **68**:529-530 (2002).
5. B.O. Palsson, "In silico biology through 'omics,'" *Nature Biotechnology*, **20**: 649-650 (2002).
6. B.O. Palsson, "In silico biotechnology: Era of reconstruction and interrogation", *Current Opinion in Biotechnology*, **15**(1): 50-51(2004).
7. Herrgard, M.J. and B.O. Palsson, "Flagellar Biosynthesis *In Silico*: Building Quantitative Models of Regulatory Networks," *Cell*, **117**: 1-20, (2004).
8. Westerhoff, H.V. and Palsson, B.Ø., "The evolution of molecular biology into systems biology," *Nature Biotechnology*, **22**(10): 1249-1252, (2004).
9. Palsson, B.Ø., "Two-Dimensional Annotation of Genomes", *Nature Biotechnology*, **22**(10): 1218-1219, (2004).
10. Fong, S.S., Joyce, A.R., and Palsson, B.Ø., "The econometrics of evolution", *Nature Chemical Biology*, **1**(4):191-192 (2005).
11. Herrgard, M.J. and Palsson, B.Ø., "Untangling the web of functional and physical interactions in yeast", *Journal of Biology*, **4**(5)(2005).

12. Oh, Y.K., Joyce A.R., and Palsson B.Ø., "Constraint-based Genome-Scale *In Silico* Models for Systems Biology" *Asia-Pacific Biotech News*, 10:123-36. (2006).

## BOOKS

1. J.R.W. Masters and B.O. Palsson Eds (1999), "Human Cell Culture, Vol 1: Cancer cell lines part 1," Kluwer Academic Publishers, Lancaster, UK.
2. J.R.W. Masters and B.O. Palsson Eds (1999), "Human Cell Culture, Vol 2: Cancer cell lines part 2," Kluwer Academic Publishers, Lancaster, UK.
3. J.R.W. Masters and B.O. Palsson Eds (1999), "Human Cell Culture, Vol 3: Cancer cell lines part 3," Kluwer Academic Publishers, Lancaster, UK.
4. M.R. Koller, B.O. Palsson, and J.R.W. Masters Eds (1999), "Human Cell Culture, Vol. 4: Primary Hematopoietic Cells," Kluwer Academic Publishers, Lancaster, UK.
5. M.R. Koller, B.O. Palsson, and J.R.W. Masters Eds (2001), "Human Cell Culture, Vol. 5: Primary Mesenchymal Cells," Kluwer Academic Publishers, Lancaster, UK.
6. B.O. Palsson, J.A. Hubbell, R. Plonsey, and J.D. Bronzino Eds (2003), "Tissue Engineering," CRC Press, Boca Raton.
7. B.O. Palsson and S.N. Bhatia, "Tissue Engineering," Prentice Hall, New Jersey, 2003
8. B.O. Palsson, "Systems Biology: properties of reconstructed networks," Cambridge University Press, 2006.
9. J.R.W. Masters and B.O. Palsson Eds (2008), "Human Cell Culture, Vol 7: Human Adult Stem Cells, Springer, The Netherlands.

## CHAPTERS IN BOOKS

1. B. O. Palsson, "Making Mathematical Descriptions of Metabolic Reaction Networks Manageable" in Activation, Metabolism and Perfusion of the Heart - Simulation and Experimental Models, Chapter 37, pp. 584-618, S. Sideman and R. Beyar, Editors, Martinus Nijhoff, Publishers, Dordrecht/Boston/Lancaster, (1987).
2. S. S. Ozturk, B. O. Palsson, J. Thiele, and G. Zeikus, "Modeling of Interspecies Hydrogen Transfer in Anaerobic Ecosystems," in Biotechnology Progress: Scaleup and Mixing, C. S. Ho and J. Y. Oldshue, Editors, pages 142-152, AIChE, New York, (1987)
3. Minoo Javanmardian and B. O. Palsson, "Design and Operation of an Algal Photobioreactor System," *Advances in Space Research* 12: 231-235, (1992).
4. J. Caldwell, B.O. Palsson, M.F. Clarke, and S.G. Emerson, "The Role of Physiological Perfusion in the Metabolism and Genetic Regulation of Cytokine Production in Mesenchymal Stromal Cells," in The Hematopoietic Microenvironment: The Functional and Structural Basis of Blood Cell Development, eds. Michael W. Long and Max S. Wicha, the Johns Hopkins University Press, Baltimore (1993).
5. J. B. Dressman, B. O. Palsson, A. G. Ozturk, and S. S. Ozturk, Chapter 11 "Mechanisms of Release from Coated Pellets," pp 285 - 306 in Multiparticulate Oral Drug Delivery, Isaac Ghebre-Sellassie, Editor, Marcel Dekker, Inc. New York 1994).
6. G.M. Lee and B.O. Palsson, "Monoclonal Antibody Production Using Free-Suspended and Entrapped Hybridoma Cells," Chapter 14 in Biotechnology and Genetic Engineering Review, vol. 12, ed. Michael P. Tombs, Intercept, Andover, Hants. UK (1994).
7. BO Palsson, MR Koller, S Rummel, J Maluta and RD Armstrong, "Expansion of Hematopoietic Stem and Progenitor Cells in Perfusion Cultures," Proceedings of the Seventh International Symposium on Autologous Marrow and Blood Transplantation, Arlington Texas, KA Dickie and A Keating Editors (1994)
8. M.R. Koller and B.O. Palsson, Tissue Engineering of Bone Marrow," Chapter 116, Section XI in The Biomedical Engineering Handbook, Editor-in-Chief Joseph D. Bronzino, CRC Press, Boca Raton, FL (1995).
9. D. Brott, M.R. Koller, S. Rummel, and B.O. Palsson, "Flow Cytometric Analysis of Cells Obtained from Human Bone Marrow Cultures," Chapter 7 in Flow Cytometry Applications in Cell Culture, M. Al-Rubeai and A.N. Emery, eds. Marcel Dekker, Inc. New York (1995)

10. SG Emerson, BO Palsson, MF Clarke, SM Silver, PT Adams, MR Koller, GV Zant, S Rummel, RD Armstrong, J Maluta, J Douville, and L Paul (1995), "In Vitro Expansion of Hematopoietic Cells for Clinical Applications," Ch 9, pgs 215-223, in "Technical and Biological Components of Marrow Transplantation," C. Dean Buckner Ed. Kluwer Academic Publishers, Boston.
11. Bernhard O. Palsson, Duk Jae Oh and Manfred R. Koller (1995) in "Animal Cell Technology: Developments Towards the 21st Century" (Beuvery, E.C., Griffiths, J.B. and Zeijlemaker, W.P., eds), pp. 1073-1081, Kluwer Academic.
12. MR Koller, RJ Maher and BO Palsson (1996), "Hydrocortisone and SCF Interact to Influence the Regulation of Hematopoietic Cell Development by the Bone Marrow Microenvironment," Keystone Symposium on Bone Marrow Transplantation.
13. Bernhard Palsson, "Hematopoietic Cells," (1997) in "Frontiers in Tissue Engineering", Eds C.W. Partrick Jr., A.G. Mikos, and L.V. McIntire, Pergamon Press
14. S. Andreadis and B.O. Palsson "Intra-cellular Stability of Retroviral Vectors: Mathematical Modeling and Kinetic Experiments," *Cancer Research, Therapy and Control*, 7:77-85, Harwood Academic Publishers
15. B.O. Palsson, "Tissue Engineering," (1998) in "Introduction to Bioengineering," eds. John Enderle, Susan M. Blanchard, and Joseph Bronzino, Academic Press.
16. M.R. Koller, B.O. Palsson, "Hematopoietic Stem and Progenitor Cells," in "Human Cell Culture, Volume 4," Kluwer Academic Publishers (2004).
17. J.S. Edwards, R. Ramakrishna, C.H. Schilling, and B.O. Palsson, "Metabolic Flux Balance Analysis," *Metabolic Engineering* 2: 13-57, (1999), S.Y. Lee, Papoutsakis, E.T., Eds; Springer-Verlag: New York.
18. J.S. Edwards, C.H. Schilling, M.W. Covert, S.J. Smith, and B.O. Palsson, "Genomic Engineering of Bacterial Metabolism," *Encyclopedia of Microbiology* 1: 32-42, (2000), Academic Press.
19. J.L. Reed, S.S. Fong, and B.O. Palsson, Phenomics, in *Microbial Diversity and Bioprospecting*, Chapter 27, p.280, A. Bull, ASM Press (2004).
20. N.D. Price, J.A. Papin, and B.O. Palsson, "In silico cells: studying genotype-phenotype relationships using constraints-based models," *Metabolic Engineering in the Post-Genomic Era*, 2003, Horizon Scientific Press.
21. Joyce A.R. and Palsson B.O. (2006) Toward whole cell modeling and simulation: Comprehensive functional genomics through the constraint-based approach, in *Systems Biological Approaches in Infectious Diseases*, 64:265-311, H.I. Boshoff and C.E. Barry, Eds (2006).
22. Hergard MJ and Palsson BO., *Genome-Scale Models of Metabolic and Regulatory Networks*, In: *Systems Biology, Volume II: Networks, Models, and Applications* Edited by Rigoutsos, I. and Stephanopoulos, G., Oxford University Press, New York, NY (2006).
23. Joyce A.R. and Palsson B.O. *In silico Genome-scale Metabolic Models: The Constraint-based Approach and its Applications*, in *Systems Biology and Synthetic Biology*, (Fu P., Latterich M., and Panke S., ed.'s), John Wiley & Brothers, Inc., Hoboken, NJ, (2007).
24. Thiele, I. and Palsson, B.O., "Bringing Genomes to Life: The Use of Genome-Scale *in silico* Models," in *Introduction to Systems Biology* (S. Choi, ed.), Humana Press, New York (2007).
25. Reed JL and Palsson BO. *Systems Biology: A Four Step Process*, in *An Introduction to Bioengineering*, edited by P.C. Chen, S. Chien, and Y.C. Fung. *World Scientific Publishing Company* (2008).
26. Joyce A.R. and Palsson B.O. Predicting Gene Essentiality Using Genome-scale *in silico* Models, in *Gene Essentiality at Genome Scale: Protocols and Bioinformatics*, (Osterman A. and Gerdes S., ed.'s). Humana Press, Totawa, NJ, (2008).
27. Lewis, N. E., Thiele, I., Jamshidi, N. and Palsson, B.O., "Metabolic Systems Biology: A Constraint-Based Approach," in *Encyclopedia of Systems Science*, (B. Meyers, ed.), Springer, New York, In Press.

## EDITORIAL ACTIVITIES

1. Special Editor, with Jeffrey A. Hubbell and E.T. Papoutsakis, of *Tissue Engineering & Cell Therapies: I and II*, special issues of *Biotechnology and Bioengineering*, 43:7 and 43:8, March 25 and April 5, 1994, John Wiley & Sons, Inc., Publishers.
2. Editorial Board, *Tissue Engineering*, Charles A. Vacanti and Antonios G. Mikos, editors, Mary Ann Liebert, Inc., Publishers. (1994 to 2000)
3. Board of Associate Editors, *Mathematical Problems in Engineering: Problems, Theories, and Applications*, V. Lakshmikantham and S.M. Meerkov, Editors, Gordon & Breach Publishing Co. (1994 to 2000)
4. Editorial Board, *Annals of Biomedical Engineering*, James B. Bassingthwaite and Daniel A. Hammer, Editors (1995 to present)
5. Section editor for *Tissue Engineering*, with Jeffrey A. Hubbell, *The Biomedical Engineering Handbook*, Editor-in-Chief Joseph D. Bronzino, CRC Press, Boca Raton (1995).
6. Editorial Board, *Biotechnology and Bioengineering*, Douglas Clark Editor-in-Chief, Wiley and Sons (1996-present)
7. Editorial Board, *Metabolic Engineering*, G. Stephanopoulos, M. Yarmush, and A. Sinskey, Editors, Academic Press.
8. Advisory Editorial Board, *Nature Molecular Systems Biology*, (2004-Present).
9. Editorial Board, *Journal of Bacteriology*, (2004-Present).
10. Editorial Board, *Journal of Biological Chemistry* (2007-Present).

## PROPOSALS ADMINISTERED

TITLE	SOURCE	AMOUNT	CO-PI/PERIOD
Mathematical Modeling of Anaerobic Digestion Dynamics	Michigan Biotechnology Institute	\$22,500	18 months, (1985).
Use of Kinetic Models to Improve Blood Storage	Whitaker Foundation	\$82,532	Nov 1985 - Oct 1987
Cell Culture Facility	University of Michigan NIH Biomedical Research Support Grant	\$59,500	Professors J. S. Schultz and H. Y. Wang (1986)
Cell Culture Facility	NSF equipment grant	\$21,300	H.Y. Wang Co-Pi (1986)
Purchase of a Spectrofluorometer	University of Michigan NIH Biomedical Research Support Grant	\$14,000	Professor J. S. Schultz (1987)
Cellular Bioengineering	From the Presidential Initiatives Fund at the University of Michigan	\$90,000/Yr.	Prof. M. Savageau (PI), with Profs M.E. Meyerhoff and A. R. Midgley, June 1987 - May 1990
Metabolic Dynamics in the Red Cell	FIRST award from NIH	\$90,000/Yr.	Sept 1987 - Aug 1992
Efficient Monoclonal Antibody Production	NSF Biotechnology Cluster Grant	\$803,520	Professors M. A. Savageau and A. R. Midgley, Sept 1987 - Sept 1990.
Life Support Systems via the Use of Cell Culture	CAMRSS, a NASA funded CCDS center	\$150,000	Prof. P. Kaufman and Dr. T. Huard, Feb 1988 - Jan 1989
Construction of a High-efficiency Ex vivo Bone Marrow	NSF Tissue Engineering Initiative	\$50,000	Drs. S. Emerson (PI) and M. Clarke in Internal Medicine, Sept 88 to Aug 89
Gravitropic Response Mechanism in Cereal Grass Shoots	NASA	\$55,000	Prof. P. Kaufman (PI) 2/89-6/89
Development of an Ex vivo Bone Marrow System	Hambrecht & Quist	\$300,000	S.G Emerson (PI) and M. Clarke 4/89 - 10/89
Development of a Photo-bioreactor and Green Plant Cell Lines for CELSS	CAMRSS, a NASA funded CCDS Center	\$250,000	Professor P. Kaufman, Feb 1989 - Sept 1989
Development of an Ex vivo Bone Marrow System	Ann Arbor Stromal Inc.,	\$900,000	S. Emerson (PI) and M. Clarke 8/89-4/92
Development of an Algal Photo-	CAMRSS, a NASA-funded	\$120,000	Nov. 1989-Sept 1990.



bioreactor	CCDS center		
Establishment of a Digital Imaging Facility	BSRG/NIH	\$27,000	Prof. J.J. Linderman (PI) April 1990 - March 1991.
Efficient Monoclonal Antibody Production	NSF Biotechnology Cluster Grant	\$664,098	Prof. M. E. Meyerhoff, Nov 1990 - Oct 1993
Photo-bioreactor Engineering: Light Delivery and Algal Growth and Gas Exchange Rates	NASA/Headquarters	\$960,713	Prof. T. L. Killeen, Jan, 1991 to Jan, 1994.
Development of an ex vivo Retroviral Infection System	Aastrom Biosciences Inc.	\$30,000	Post-doctoral Fellowship 1/14/91 - 10/13/92.
Development of a Hematopoietic Bioreactor System	Aastrom Biosciences Inc.	\$45,000	Post-doctoral Fellowship 5/1/92 - 10/31/93.
Ex vivo Growth and Manipulation of Human Hematopoietic Cells - Hematopoietic Cell Expansion Bioreactor Design and Retroviral Gene Transfer	Aastrom Biosciences Inc.	\$532,259	3/1/92 - 6/30/93.
Shear Sensitivities of Human Bone Marrow Cultures	NASA Headquarters	\$315,000	12/1/92-12/1/93.
Hematopoietic Bioreactor Design with Applications to Gene Therapy	Aastrom Biosciences, Inc.	\$344,001	7/1/93 - 12/31/94.
Biological Determinants of Photobioreactor Design	Department of Energy	\$99,898	9/1/93 - 8/31/95.
Hematopoietic Cell Expansion System	NIH SBIR Phase II Award	\$520,000	3/1/95 - 2/29/96.
Biochemical Engineering: Genomatics and Whole Cell Simulators	UC Biotechnology Program	\$80,000	7/1/96-6/30/98
Shear Sensitivities of Human Bone Marrow Cultures	NASA Biotechnology Program	\$190,000	10/1/95-12/1/97.
Stem Cell Motility	The Stern Foundation	\$50,000	11/1/97-10/30/98
Hematopoietic Stem Cell Motility	National Institutes of Health, R01 HL59234	\$152,207 annual direct costs	Jan 1, 1998 to Dec 31 2001
Genomically Based Models for Antibiotic Development	National Institutes of Health, R01 GM57089	\$123,000 annual direct costs	7/1/98 - 6/30/01
Mechanisms of Stem Cell Migration	National Institutes of Health, R01 HL 60398	\$200,000 annual direct costs	7/10/98 to 7/9/02
In Silico Analysis of the Escherichia Coli Metabolic Genotype and the Construction of Selected Isogenic Strains	National Science Foundation/Department of Defense, BES-9814092	\$88,388/Yr.	3/1/99-2/28/02
Computational Infrastructure for Engineering Microorganisms	National Science Foundation/KDI, SBR-9873384	\$297,218	10/1/98-9/30/00
Genome-scale <i>in silico</i> Model for <i>E. coli</i>	National Institutes of Health, R01 GM057089	\$250,000	8/1/98-4/30/11
Tissue Engineering	Whitaker Foundation/Teaching Materials Project	\$89,962	11/1/99-1/1/03
Kinetic and Regulatory Constraints on Metabolism	National Science Foundation/BES-0120363	\$469,934	9/1/01-8/31/03
Testable <i>in silico</i> Hypotheses for <i>E. coli</i> Growth	National Institutes of Health, R01 GM62791	\$432,254 annual direct costs	4/01/01-5/31/10
Genomically Based Models for Antibiotic Development (Renewal)	National Institutes of Health, R01 GM57089	\$250,000 annual direct costs	4/1/03 - 3/31/08
Systems Biology and Bioengineering	Whitaker Foundation/Teaching Materials Project	\$84,713	4/1/03-11/30/04
Network Based Analysis of Kinetics and Regulation	National Institutes of Health/R01 GM68837	\$225,000 annual direct costs	7/15/03-6/30/07
Reconstruction and Simulation of	National Science	\$432,697	10/1/03-9/30/06

Genome-Scale Regulatory Networks	Foundation/BES-0331342		
A Genome-Scale Regulated Metabolic Model of Yeast	National Institutes of Health, RO1 GM071808	\$446,123 annual direct costs	7/01/04-6/30/09
Analysis of the Genetic Potential and Gene Expression of Microbial Communities Involved in the in situ Bioremediation of Uranium and Harvesting Electrical Energy from Organic Matter	DOE, 03-003721 G 00	\$254,750	9/1/06-1/31/09
Systems-Level Understanding of Hydrogen Production by <i>Thermotoga maritima</i> .	DOE, DE-PS02-08ER08-12	\$506,489	9/15/08-9/14/11
A Systems Biology Program to Study Infectious Microorganisms in Taiwan	NHRI (Taiwan)	\$1,052,904	7/1/08-8/31/11

#### INVITED TALKS PRESENTED AT MEETINGS:

1. B. O. Palsson "Making Mathematical Descriptions of Metabolic Reaction Networks Manageable" Presented at the *3rd Henry Goldberg Workshop* on "Simulation and Modeling of the Cardiac System: from Cellular Activation to Muscular Activity", March 31 - April 2, 1986, Rutgers University, New Brunswick, NJ.
2. O. Palsson, A. Joshi, and S. S. Ozturk, "Reducing Complexity in Metabolic Networks: Making Metabolic Meshes Manageable", *FASEB meeting*, St. Louis, MO, April 14-18, 1986.
3. B. O. Palsson, I-der Lee, and A. Joshi, "A Comprehensive Computer Model of Human Red Cell Metabolism," *International Symposium on Mathematical Models of Cellular Processes*, Holzhau, GDR, November 19-23, 1989
4. B. O. Palsson, S. Emerson, M. Clarke, R. Schwartz, and J. Caldwell, "Reconstitution of a Functioning Bone Marrow," *The 1989 International Chemical Congress of Pacific Basin Societies*, Symposium on Cell Culturing, Honolulu, Hawaii, December 17-22, 1989
5. B. O. Palsson, "Metabolic Modeling, Design and Operation of Continuous Perfusion Hematopoietic Cultures," *UCLA Symposium on Tissue Engineering*, Keystone CO, April 5-12, 1990
6. Mino Javanmardian and B. O. Palsson, "Can Biotechnology Help with Global Warming?," *American Chemical Society 200th National Meeting*, Washington, D.C., August 26-31, 1990
7. J. D. Keasling, B. O. Palsson, and S. Cooper, "Cell-cycle-specific F' lac Plasmid Replication: Regulation by Cell Size Control of Initiation," *European Molecular Biology Organization Meeting on the Bacterial Cell Cycle*, Collouges-LaRouge, France, September 30 - October 5, 1990.
8. B. O. Palsson, "Hematopoietic Bioreactor Systems," Engineering Foundation Conference - *Cell Culture Engineering III: Tissue Engineering*, Palm Coast, Florida, February 2 - 7, 1992.
9. A. Peng and C. G. Lee, "Two Novel Bioreactor Systems for the Cultivation of Human Bone Marrow and for Growing Algae Photoautotrophically," *SIM/CSM Joint Meeting - Symposia on Novel Bioreactors*, Toronto, Canada, August 1 - 6, 1993.
10. B. O. Palsson, "Hematopoietic Tissue Engineering: From Basic Principles to Clinical Practice," *IFMBE 1st International Conference on Cellular Engineering*, Stoke-on-Trent, UK, September 12 - 15, 1993.
11. B. O. Palsson, "Cell and Tissue Engineering: An Emerging Discipline?" *N + N Meeting on Cellular Engineering*, USA and UK, Chester, UK, September 15-17, 1993.
12. B. O. Palsson, "Tissue Engineering and Differentiating Cell Types," *Frontiers in Bioprocessing III*, Boulder, CO, September 19-23, 1993.
13. B. O. Palsson, "Shear Sensitivities of Human Bone Marrow Culture," *Investigators' Meeting*, NASA/Johnson Space Center, Houston, TX October 15-16, 1993.
14. C. G. Lee and B. O. Palsson, "Design and Performance of an LED-Based Algal Photobioreactor," *International Winter Meeting of the American Society of Agricultural Engineers (ASAE)* Chicago, IL, December 14-17, 1993.

15. B. O. Palsson, "Hematopoietic Perfusion Bioreactors: Scientific and Clinical Utility," Keystone Symposia on Molecular and Cellular Biology: *Tissue Engineering*, Taos, New Mexico, February 20-26, 1994
16. B. O. Palsson, "Expansion of Progenitor Cells," *The Seventh International Symposium on Autologous Bone Marrow Transplantation*, Arlington, TX, August 17-20, 1994.
17. B. O. Palsson, "Hematopoietic Tissue Engineering," *ESACT/IAACT Meeting Animal Cell Technology 'Developments towards the 21st Century'*, Veldhoven, The Netherlands, September 12-16, 1994.
18. B. O. Palsson, "Perfusion-Based Bioreactors for the Expansion of Human Bone Marrow, Stem and Progenitor Cells," *BPEC Symposium and Workshop on Bioprocessing Needs for Cell Based Therapies*, MIT, Cambridge, MA, January 18-19, 1995.
19. B. O. Palsson, "Tissue Engineering of Bone Marrow," *National Engineering Forum*, UC San Diego School of Engineering, San Diego, May 4-5, 1995.
20. B. O. Palsson, M.R. Koller and J. Maluta, "Hematopoietic Tissue Engineering Challenges in Producing Clinical Scale Cell Populations," *Engineering Foundation Biochemical Engineering IX: Interdisciplinary Foundations for Creating New Biotechnology*, Davos, Switzerland, May 21-26, 1995.
21. B. O. Palsson, M.R. Koller and J. Maluta, "Hematopoietic Tissue Engineering Challenges in Producing Clinical Scale Cell Populations," *2nd International Meeting on Cell Engineering*, San Diego, August 1995
22. B. O. Palsson and A. Varma, "Metabolic Flux Balancing: basic concepts, scientific and practical use," *Recent Advances in Fermentation Technology*, San Diego, November 4-7 1995
23. BO Palsson, J. Maluta, CA Peng, RD Armstrong and MR Koller, "Scientific and Clinical Applications of and Ex vivo Hematopoietic Model," *Tissue Engineering*, Keystone Symposia, Taos, NM Jan 1996
24. BO Palsson, AS Chuck and G Huang, "Gene Therapy: The Importance of Random Brownian Motion," *Cell Culture Engineering V*, Engineering Foundation Meetings, San Diego, Jan 1996
25. BO Palsson, AS Chuck and G Huang, "Retroviral Infection is Limited by Random Brownian Motion," *Gene Therapy for Hematopoietic Stem Cells in Genetic Disease and Cancer*, Keystone Symposia, Taos, NM Feb 1996
26. BO Palsson, "Retrovirally-Mediated Gene Transfer is Limited by Random Brownian Motion," 4th International Symposium on Recent Advances in Hematopoietic Stem Cell Transplantation: *Clinical Progress, New Technology, and Gene Therapy*. UCSD/ISHAGE, San Diego, CA, April 11-13, 1996
27. BO Palsson, "How Should We Approach the 'Engineering' of Metabolic Function," *Recombinant DNA Biotechnology: Focus on Metabolic Engineering*, Engineering Foundation Conferences, Fercroft Conference Resort, Danvers, MA, October 6-11, 1996
28. BO Palsson, "Modeling Challenges in Tissue Engineering and Complex Systems," *Modeling in Biochemical Engineering*, October 11-12, 1996 at the University of Minnesota, Minneapolis/St. Paul, MN.
29. BO Palsson, "Genetic Circuits," *Chemical Engineering and Living Systems*, in Honor of Professor EN Lightfoot, Nov 9-10, 1996, Madison, WI
30. J Edwards, R Ramakrishna and BO Palsson, "Significant Flexibility Exists in the Core Metabolic Pathways," Soc. for Industrial Microbiology Annual Meeting, Reno, NV, Aug 3-7th, 1997
31. B.O. Palsson "What lies Beyond Bioinformatics?," National Academy of Engineering, Annual Meeting, October 6 to 8, 1997, Washington DC
32. B. O. Palsson "The Importance of Stem Cells in Tissue Engineering: lessons learned from hematopoiesis," *Tissue Engineering*, Keystone Symposia, Copper Mountain, Jan 1998
33. B. O. Palsson "The Importance of Stem Cells in Tissue Engineering: lessons learned from hematopoiesis," *Cell Culture Engineering VI*, Engineering Foundation Conference, Pacific Beach, February 8-14, 1998
34. BO Palsson, "What Lies Beyond Bioinformatics?," Bioinformatics Workshop, Sat. April 4, 1998 San Diego Supercomputer Center Auditorium
35. BO. Palsson, "Mechanisms of Stem Cell Migration," 6th International Symposium on Recent Advances in Hematopoietic Stem Cell Transplantation," April 16-18 San Diego, 1998
36. BO. Palsson, "Bioinformatics—Their Role in Modern Scientific Investigation and Cardiac Disease," 6th Antwerp-La Jolla-Kyoto Research Conference on Cardiac Function, La Jolla, April 25-27, 1998

37. B.O. Palsson, "Towards Metabolic Phenomics: analysis of genomic data using flux balances," Metabolic Engineering II, Elmau, Germany, Oct 25-30, 1998
38. B.O. Palsson, "Haemophilus influenzae Metabolic Genotype; Its definition, Systems, Characteristics, and Capabilities," AIChE Annual Meeting, Miami Beach, FL November, 1998.
39. B.O. Palsson, "Synthesizing and Characterizing *In Silico* Bacterial Strains," *In Silico* Biology Meeting, San Francisco, CA, June 1999.
40. B.O. Palsson, "Building Metabolic Models from Annotated Genes," Biochemical Engineering XI: Molecular Diversity In Discovery and Bioprocessing, Park City, UT, July 1999.
41. B.O. Palsson, "Life on the Edge: using genome-scale *in silico* models of microorganisms to interpret and predict metabolic phenotypes," 11th International Genome Sequencing and Analysis Conference, Miami, FL, September 1999.
42. B.O. Palsson, "What Lies Beyond Bioinformatics?," BMES/EMBS Joint Meeting, Atlanta, GA, October 1999.
43. B.O. Palsson, J.S. Edwards, and R.I. Ibarra, "The generation of experimentally testable hypotheses *in silico* using an *Escherichia coli* metabolic model, 4<sup>th</sup> Annual Hilton Head Workshop, Hilton Head, SC, February, 2000.
44. B.O. Palsson, "Reconstruction of Metabolic Networks *in silico* and Formulation of Testable Experimental Hypotheses," Bioinformatics 2000, Elsinore, Denmark, April, 2000.
45. B.O. Palsson, "Living on the edge: *E. coli* optimizes its growth within governing physico-chemical constraints," *In Silico* Biology Conference, San Francisco, June, 2000.
46. B.O. Palsson, "Functional Genomics," The Whitaker Foundation Biomedical Engineering Education Summit, Arlington, VA, December, 2000
47. B.O. Palsson, "Models of microorganisms to interpret and predict metabolic phenotypes: basic concepts, scientific and applied uses," NASCRE 1 Conference, Houston, TX January, 2001.
48. B.O. Palsson, "Single and multi cellular communication pathways and their transduction," DOE/NSF Workshop on Biological Information Processing and Systems," Greenville, SC, January, 2001.
49. B.O. Palsson, "The challenges of *in silico* biology," Workshop on Challenges and Opportunities in Data Management, Palo Alto, CA, March, 2001.
50. B.O. Palsson, "Life on the edge: Using genome-scale *in silico* models of microorganisms to interpret and predict metabolic phenotypes," Recovery of Biological Products 10, Cancun, Mexico, June, 2001
51. B.O. Palsson, "Influence of Bioinformatics on Metabolic Engineering," CAB8 Conference, Quebec, Canada, June, 2001.
52. B.O. Palsson, "Do *in silico* models of metabolism represent their *in vivo* counterparts well?" Beyond Genome 2001, San Francisco, June, 2001.
53. B.O. Palsson, "The Phase transition from *in vivo* to *in silico* biology," The Bayer Lectures in Chemical Engineering, Berkeley, September, 2001.
54. B.O. Palsson, "Life on the edge: using genome scale *in silico* models of microorganisms to interpret and predict metabolic phenotypes," DOE 9<sup>th</sup> International Conference on Microbial Genomes, Gatlinburg, TN, October, 2001.
55. B.O. Palsson, "Data-Driven Constraints-Based Models in Biology," AIChE Meeting, Reno, NV, November, 2001.
56. B.O. Palsson, "The Phase transition from *in vivo* to *in silico* biology," Syngenta, December, 2001
57. B.O. Palsson, "The Phase transition from *in vivo* to *in silico* biology," CHI Metabolomics Meeting, December, 2001
58. B.O. Palsson, "The phase transition from *in vivo* to *in silico* biology," ETH, Zurich, Switzerland, January, 2002.
59. B.O. Palsson, "The phase transition from *in vivo* to *in silico* biology," Novartis, Basel Switzerland, January, 2002.
60. B.O. Palsson, "Data Driven Constraint-Based Models in Cell Biology," 6<sup>th</sup> Annual Lake Tahoe Symposium, Granlibakken, Lake Tahoe, CA, January, 2002.
61. B.O. Palsson, "From bioreactor design to genetic circuits," Lindbergh-Carrel Symposium, Charleston, SC, February, 2002.
62. B.O. Palsson, "Model-Centric Integrated Genomic Databases," Frontiers of Genomics, Madison, Wisconsin, May 16-17, 2002.
63. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models" Cuernavaca, Mexico, May 23-25, 2002.

64. B.O. Palsson, "The New Biotechnology Seminar Series," Campbell and Flores, LLP, June, 2002.
65. B.O. Palsson, "Keynote Speaker," IEEE Computer Society Bioinformatics Conference, Stanford University, August 14-16, 2002.
66. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models," Engineering Foundation Conferences, Metabolic Engineering IV, Barga, Italy, October 6-11, 2002.
67. B.O. Palsson, FPB Division Award Talk, "The phase transition from *in vivo* to *in silico* biology," AIChE Meeting, Indianapolis, IN, November 3-5, 2002.
68. B.O. Palsson, "Genome-Scale Models for Prospective Metabolic Engineering," CHI—Metabolic Engineering: Modifying Metabolic Pathways Conference, Raleigh, NC, December 2-3, 2002.
69. B.O. Palsson, "Laser-Enabled High-Throughput High-Content Single-Cell Analysis," IBC—Cell Based Assay and Screening, Philadelphia, PA December 4-6, 2002.
70. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models," Harvard University Department of Genetics, December 10, 2002.
71. B.O. Palsson, "Computer-assisted search for new antimicrobials," 2002 CSPA ANTI Special Session, Ft. Lauderdale, FL, December, 2003.
72. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models, ICSB 2002, Stockholm, Sweden, December 13-15, 2002.
73. B.O. Palsson, "A model-driven analysis of expression data for *E. coli* and for Yeast," EMBO Practical Course, University of Milano-Bicocca, Italy, January 13, 2003.
74. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models," Plenary Talk—Enzymes & Biocatalysis for Drug Discovery and Development, San Diego, CA, January 30, 2003.
75. B.O. Palsson, "Genome-scale analysis for new metabolic engineering procedures," National Science Foundation Metabolic Engineering Conference, January 31, 2003.
76. B.O. Palsson, "Bringing Genomes to Life: The Key Role of Genome-scale Computer Models," IRI Frontiers of Technology 2003 Conference, February 27-28, 2003.
77. B.O. Palsson, "Novel HTC?HTS Technology," BiolT Meeting, Boston, March 26, 2003.
78. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-scale *in silico* Models," NIH Computational Approaches to Biological Systems Seminar Series, March 27, 2003.
79. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-scale *in silico* Models," IBM Research Computational Biology Seminar, March 28, 2003.
80. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-scale *in silico* Models," UCSD Bioinformatics Symposium, March 29, 2003.
81. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-scale *in silico* Models," Plenary Talk—Northwestern University Computational Science and Engineering Spring Symposium, April 3, 2003.
82. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models" Duke Center for Bioinformatics and Computational Biology Conference, May, 2003.
83. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models," Keynote Address—ESCAT Meeting, Spain, May, 2003.
84. B.O. Palsson, "Chairperson's Remarks," CHI Beyond Genome 2003, San Diego, June, 2003.
85. B.O. Palsson, "E. coli i2K," IECA 2003, Tsuruoka, Japan, June 21-26, 2003.
86. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models," The Nyhan Center Planning Meeting, Palo Alto, September, 2003.
87. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models," GTL Data Standards Workshop, San Francisco, CA, September, 2003.
88. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models," ERATO Kitano Project, Tokyo, Japan, September, 2003.
89. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models," Southern California Biotechnology Symposium, Laguna Beach, CA, September, 2003.
90. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models," NAE Annual Meeting, Washington, DC, October, 2003.
91. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models," International Symposium on New Horizons in Molecular Sciences and Systems: An Integrated Approach, Okinawa, Japan, October, 2003.
92. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models," IBC Systems Biology for Drug Discovery and Development, Boston, MA, October, 2003.
93. B.O. Palsson, "Bringing Genomes to Life: The Use of *in silico* Models," NIH SysBio SIG Retreat, Washington, DC, November, 2003.

94. B.O. Palsson, "Systems Biology and Genetic Circuits," 2004 Gordon Conference in Molecular Evolution, Ventura, CA, February, 2004.
95. B.O. Palsson, "Metabolomics," Mass Spectrometry in Systems Biology, Keynote Address, Santa Fe, NM, February, 2004.
96. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-Scale *in silico* Models," Rutgers Collaboratus, XIV, Rutgers University, March, 2004.
97. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-Scale *in silico* Models," Invited Seminar, Caltech, April, 2004.
98. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-Scale *in silico* Models," ISAC International Congress, Plenary Talk Montpellier, France, May, 2004.
99. B.O. Palsson, "E. Coli 2K: A Chemically and genetically structured framework for relating the virtual and real Gemini," Second International E. Coli Alliance Conference on systems Biology, Alberta, Canada, June, 2004.
100. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-Scale *in silico* Models," the 250<sup>th</sup> Anniversary of Columbia University—A Symposium for Biomedical Engineering, New York, September, 2004.
101. B.O. Palsson, "Combining Genome-Scale Reconstructions, OptKnock and Adaptive Evolution to Design Production Strains," Metabolic Engineering V, Lake Tahoe, September, 2004.
102. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-Scale *in silico* Models," The University of Texas, Austin Seminar Series, Austin, September, 2004.
103. B.O. Palsson, "Critical Issues in Systems Biology," IBC Systems Biology Summit, Boston, September, 2004.
104. B.O. Palsson, "Using *in silico* Models of Microorganisms to Elucidate Metabolic Pathways," Arrowhead 2004, Lake Arrowhead, September, 2004.
105. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-Scale *in silico* Models," BioJapan 2004, Tokyo, September, 2004.
106. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-Scale *in silico* Models," 2<sup>nd</sup> International Conference on Pathways, Networks, and Systems, Crete, October, 2004.
107. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-Scale *in silico* Models," Conference on Complex Systems, Northwestern University, Evanston, October, 2004.
108. B.O. Palsson, "Genome-scale Assessment of Phenotypic Changes During Adaptive Evolution," Keynote Address, Heidelberg, October, 2004.
109. B.O. Palsson, "Systems Biology: Bringing Genomes to Life," Neural Information Processing Conference 2004, Vancouver, Canada, December, 2004.
110. B.O. Palsson, "Educational and research challenges in systems biology," CSUPERB Symposium, Los Angeles, January, 2005.
111. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-Scale *in silico* Models," Keystone Symposium, Santa Fe, February, 2005.
112. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-Scale *in silico* Models," Stanford University BIO-X Lecture, March, 2005.
113. B.O. Palsson, "Bringing Genomes to Life: The Use of Genome-Scale *in silico* Models," UC Berkeley, March, 2005.
114. B.O. Palsson, "New 'Dimensions' in Genome Annotation," Keystone Symposium, Denver, CO, April, 2005.
115. B.O. Palsson, "New 'Dimensions' in Genome Annotation," Institute for Systems Biology 4<sup>th</sup> Annual International Symposium, Seattle WA, April, 2005.
116. B.O. Palsson, "New 'Dimensions' in Genome Annotation," Keystone Symposium, Denver, CO, April, 2005.
117. B.O. Palsson, "New 'Dimensions' in Genome Annotation," Lewis Sigler Institute Seminar, Princeton University, April, 2005.
118. B.O. Palsson, "The plasticity of the E. coli metabolome during adaptive evolution," Metabolomics 2005, Tokyo, June, 2005.
119. B.O. Palsson, "New 'Dimensions' in Genome Annotation," SIAM Symposium, New Orleans, July, 2005.
120. B.O. Palsson, "New 'Dimensions' in Genome Annotation," BioScience 2005, Glasgow, July, 2005.
121. B.O. Palsson, "Systems Biology and Bioengineering Curriculum at UCSD," FOSBE Conference, Santa Barbara, August, 2005.

122. B.O. Palsson, "New 'Dimensions' in Genome Annotation," 12<sup>th</sup> European Congress on Biotechnology, Copenhagen, August, 2005.
123. B.O. Palsson, "Genome Scale Analysis of *E. coli* Adaptation to New Growth Environments, Including Re-Sequencing and Expression Profiling," 13<sup>th</sup> Annual International Conference on Microbial Genomes, Madison, Wisconsin, September, 2005.
124. B.O. Palsson, "New 'Dimensions' in Genome Annotation," CCB/MaRS Research Symposium, Toronto, September, 2005.
125. B.O. Palsson, "New 'Dimensions' in Genome Annotation," University of Illinois special seminar, Urbana-Champaign, October, 2005.
126. B.O. Palsson, "The construction of an integrated metabolic and transcriptional regulatory network," ICSB Yeast Systems Biology Workshop, October, 2005.
127. B.O. Palsson, "New 'Dimensions' in Genome Annotation," YSBN Workshop, Boston, October, 2005.
128. B.O. Palsson, "New 'Dimensions' in Genome Annotation," Special Seminar at EPFL, Switzerland, November, 2005.
129. B.O. Palsson, "Prospect for Cell-Scale *in silico* Models of Production Cell Lines," "Cell Line Development and Engineering Conference, St. Louis, November 2005.
130. B.O. Palsson, "New 'Dimensions' in Genome Annotation, Pacificchem Conference, Honolulu, December, 2005.
131. B.O. Palsson, "New 'Dimension' in Genome Annotation," O.A. Hougen Symposium, Madison, Wisconsin, February, 2006.
132. B.O. Palsson, "Genome-Scale Analysis of *E. coli* Adaptation to New Growth Environments, Including Re-Sequencing, Location Analysis and Expression Profiling," Systems Biology: Integrating Biology, Technology and Computation, Taos, New Mexico, March, 2006.
133. B.O. Palsson, "Systems Biology," University of Iceland, Reykjavik, March, 2006.
134. B.O. Palsson, "New 'Dimensions' in Genome Annotation," UCSF, May, 2006.
135. B.O. Palsson, "Uses of the NimbleGen oligo-arrays for re-sequencing bacterial genomes and for mapping their genome-scale transcriptional regulatory networks," University of Iceland, Reykjavik, May, 2006.
136. B.O. Palsson, "Metabolic Characterization of Bacterial Adaptation," The Second Scientific Meeting of the Metabolomics Society, Boston, June, 2006.
137. B.O. Palsson, "Reconstruction of the global human metabolic networks based on sequence annotation and legacy data," Amylin Pharmaceuticals Research Seminar, La Jolla, July, 2006.
138. B.O. Palsson, "Reconstruction of the genome-scale human metabolic network based on build-35 and bibliomic data," 2006 Synthetic Biology Symposium, Irvine, September, 2006.
139. B.O. Palsson, "," Geobacter Conference, University of Massachusetts, October, 2006.
140. B.O. Palsson, "Bottom-up reconstruction of the human metabolic network based on build-35 and bibliomic data," ICSB, Yokohama, October, 2006.
141. B.O. Palsson, "The genetic basis for adaptation of *E. coli* to new growth environments," **Plenary Lecture**, IECA Conference, Jeju, Korea, October, 2006.
142. B.O. Palsson, "Systems biology and the reconstruction of the genome-scale human metabolic network," ACDP 2006 Annual Meeting, Costa Rica, November, 2006.
143. B.O. Palsson, "The genetic basis for adaptation," **Special Lecture**, Center for Plant Cell Biology, UCR, February, 2007.
144. B.O. Palsson, "The challenge of incorporating regulatory effect in genome scale networks," Metabolic Engineering Workshop, Bethesda, February, 2007.
145. B.O. Palsson, "New 'dimensions' in genome annotation," Bioinformatics: Building Bridges Symposium, University of Minnesota, April, 2007.
146. B.O. Palsson, "Genome-scale reconstruction of the human metabolic network: basic concepts and practical uses," Nutrigenomics and Beyond: Informing the Future, "National Academy of Sciences, Washington, DC, May, 2007.
147. B.O. Palsson, "Metabolic Modeling," Summit on Systems Biology, Virginia Commonwealth University, June, 2007.
148. B.O. Palsson, "Systems Biology and Metabolomics: Mapping the Human Metabolome," Metabolomics Society 3<sup>rd</sup> Annual Conference, Manchester, UK, June, 2007.
149. B.O. Palsson, "Reconstruction of the genome-scale transcriptional regulatory network in *E. coli*," First Q-bio Conference on Cellular Information Processing, Santa FE, August, 2007.

150. B.O. Palsson, "Bringing Microbial Genomes to Life: the use of genome-scale *in silico* models for discovery and development, BioKorea 2007, Seoul, September, 2007.
151. B.O. Palsson, "Systems Biology of SNPs," 7<sup>th</sup> International Workshop on Pharmacodynamics of AntiCancer Agents, Costa Rica, September, 2007.
152. B.O. Palsson, "Genome-scale human metabolic reconstruction: current and future applications, BMES 2007, Hollywood, September, 2007.
153. B.O. Palsson, "Reconstruction of the genome-scale transcriptional regulatory network in *E. coli*," Seminar at the University of Washington, October, 2007.
154. B.O. Palsson, "New Dimensions in Genome Annotation," The Richard S.H. Mah Lecture, Northwestern University, Evanston, October, 2007.
155. B.O. Palsson, "The Reconstruction of the Genome-scale Transcriptional Regulatory Network in *E. coli*," IGSB Seminar, University of Chicago, May 2008.
156. B.O. Palsson, "Reconstruction of the Genome-scale Metabolic and Transcriptional Regulatory Networks in *E. coli*," ASM General Meeting, Boston, June, 2008.
157. B.O. Palsson, "The Genetic basis for Adaptive Evolution in *E. coli*," 16<sup>th</sup> International Microbial Genomes Conference, Lake Arrowhead, September, 2008.
158. B.O. Palsson, "The Use of Genome Scale Models for Metabolic Engineering," **Plenary Lecture**, Metabolic Engineering VII, Puerto Vallarta, September, 2008.
159. B.O. Palsson, "Global Reconstruction of the Human Metabolic Network based on Genomic and Bibliomic Data," **Keynote Address**, IBC Discovery 2 Diagnostics Conference, San Diego, October, 2008.
160. B.O. Palsson, "Reconstruction of the Genome-Scale Human Metabolic map Based on Build-35 and Bibliomic Data," **The Ernst W. Bertner Memorial Award Lecture**, Systems Biology of Cancer, Houston, October, 2008.
161. B.O. Palsson, "Systems Biology: An Era of Reconstruction and Interrogation, AiChE Annual Meeting, Philadelphia, November, 2008.
162. B.O. Palsson, "Reconstruction of the genome-scale human metabolic network based on build-35 and bibliomic data," ICB Conference on Biomolecular Engineering, Santa Barbara, CA January, 2009.

## GRADUATE STUDENTS & POST DOCTORAL RESEARCHERS

LAST	FIRST	YEAR/DEGREE	RESEARCH PROJECT	CURRENT POSITION
Applebee	Kenyon	Current		UCSD Doctoral Student
Allen	Timothy	2005/Ph.D.	Genome-scale integrated models	University of Virginia
Andreadis	Stylianos	1996/Ph.D.	Dynamics of Retrovirus Mediated Gene Transfer	SUNY, Buffalo, New York
Asthaigiri	Dilip	1994/MS	Systemic Behavior of Growth Factor Networks	University of Delaware
Barret	Christian	Current/POST		UCSD Post Doctoral Researcher
Becker	Scott	Current	Reconstruction and analysis of biochemical reaction networks	UCSD Ph.D. Student
Bell	Steve	2004-PGR		UCSD Post Doctoral Researcher
Cho	Byung-Kwan	Current/POST		UCSD Post Doctoral Researcher
Charusanti	Pep	Current/Post		UCSD Post Doctoral Researcher
Chuck	Alice	1995/Ph.D.	Directed Retroviral Motion as a Means of Enhancing Gene Transfer for Gene Therapy	Amgen, Thousand Oaks, CA—Research Scientist
Covert	Markus	2003/PhD	Metabolic Regulation in <i>in silico</i> in <i>E. coli</i>	Post Doc, Cal Tech
Duarte	Natale	2006/Ph.D.	Reconstruction and analysis of eukaryotic metabolic networks	Exponent



Edwards	Jeremy	1999/ Ph.D.	Functional Genomics and the Computational Analysis of Bacterial Metabolism	University of New Mexico
Elkins	James	Current/ POST		UCSD Post Doctoral Researcher
Fahland	Tom	1999- 2000-- PGR	Whole cell simulation of kinetics	Researcher- Genomatica
Famili	Iman	2003/ PhD	In silico Analysis of Metabolism In the single cellular Eukaryote, <i>Saccharomyces cerevisiae</i>	Researcher- Genomatica, Inc.
Feist	Adam	Current		UCSD Graduate Student
Fong	Stephen	2004/ PhD	Gene knockouts and adaptive evolution	Virginia Commonwealth University
Francis	Karl	1999/ Ph.D.	An Ex Vivo Study of the Cellular Fate Process of Migration in Immature Hematopoietic Cells	Researcher-Biocyte
Fu	Patrick	1999- 2001- PGR	Expression analysis and growth experiments in yeast	University of Hawaii
Halberstadt	Craig	1991/ Ph.D.	Transtubular Bioreactor Design	Carolinas Medical Center, Carolinas Med. Center
Herring	Chris	2006 POST	Gene regulatory network in <i>E. coli</i>	Researcher, Mascoma
Herrgard	Markus	2004/ PhD	Regulatory network reconstruction and modeling	UCSD Post-Doc
Hua	Qiang	2004/ POST	Fluxomics analysis of microorganisms	UCSD Post-Doc
Javanmardian	Minoo	1991/ Ph.D.	High Density Photoautotrophic Algal Cultures: Photobioreactor Design and Cell Cycle Kinetics	Booz Allen Hamilton
Jamshidi	Neema	2002/M S-Bioinf.	Red Blood cell simulation	UCSD PhD Student
Joshi	Abhay	1988/ Ph.D.	Integrated Metabolic Dynamics in the Human Red Cell	Sr. VP, CTO, CoTherix
Joyce	Andrew	2007/Ph D		Philip Morris
Keasling	Jay	1991/ Ph.D.	Dynamics and Control of Bacterial Plasmid Replication	UC Berkeley, Berkeley, CA--Associate Professor
Kim	Byung-Soo	1995/ MS	Bone Marrow Cell Culture and Metabolism in Rotating Wall Vessels	Professor-Korea
Lee	Baekseok	2003- 2005		Post Doctoral Researcher--Inha University
Lee	Gyun-Min	1990/ Ph.D.	Production of Monoclonal Antibodies Using Free and Immobilized Hybridoma Cells	Korean Advanced Institute of Science and Technology
Lee	Choul-Gyun	1994/ Ph.D.	Photobioreactor Engineering: High-Density Algal Cultures Using Light-Emitting Diodes	Inha University, Korea--Asst. Professor
Lee	I-Der	1992/ POST	Computer Model of Red Cell Metabolism	MERCCK--Sr. Res. Scientist
Lewis	Nathan	Current		UCSD Doctoral Student
Levee	Minette	1993/ MS	Encapsulated Bone Marrow Cultures as a Potential Assay for Human Hematopoietic Progenitors	
Mandalam	Ramkumar	1994/ Ph.D.	Nutritional Effects in High Density <i>Chlorella Vulgaris</i> Cultures	Cellerant Therapeutics
Marciniak	Jennifer	2002/M S	Bacterial evolution	UCSD Ph.D. Student
Merritt	Stephen	1992/ MS	MRNA Levels for Heavy and Light Chains in Stressed Hybridoma Cells	
Mo	Monica	Current		UCSD PhD student
Oh	Duk Jae	2000/ POST	Development of Bioreactors for Bone Marrow Cell Culture	Korean Advanced Institute of Science and Technology
Ozturk	Asuman	1990/ MS	Drug Dissolution Across Ethylcellulose Matrices	CHIRON, Emoryville, CA--Research Scientist

Ozturk	Sadettin	1990/ Ph.D.	Kinetic Characterization of Hybridoma Growth and Monoclonal Antibody Production Rates	Medimmune, Inc
Paek	Se-Hwan	1992/ POST	Continuous Bone Marrow Perfusion Chamber	Korean Advanced Institute of Science and Technology
Papin	Jason	2004/ PhD	Signaling networks	Asst. Prof—University of Virginia
Price	Nathan	2005/ PhD	Single cell models	Post Doctoral Fellow—Institute for Systems Biology
Peng	Ching-An	1995/ Ph.D.	Hematopoietic Tissue Engineering: Effects of Fluid Dynamics and Mass Transfer on Cell Growth	USC, Los Angeles, CA—Asst. Professor
Portnoy	Vasily	Current		UCSD Doctoral Student
Ramakrishna	Ramprasad	1999/ POST	Genomatics and Antibiotic Development	Physiome Sciences, New Haven, CT
Raghuathan	Anu	2002- 2005	Adaptive evolution	Mt. Sinai Medical Center
Reed	Jennie	2006/Ph D	E. coli K12	Assistant Professor, UW, Madison
Resendis-Antonio	Osbaldo	2005- 2007		UNAM
Savinell	Joanne	1991/ Ph.D.	A Quantitative Study of Hybridoma Metabolism and Monoclonal Antibody Production	Cleveland State University, Cleveland, OH—Asst. Professor
Scheilenberger	Jan	Current		UCSD Doctoral Student
Schilling	Christophe	2000/ Ph.D.	On Systems Biology and the Pathway Analysis of Metabolic Networks	Genomatica, Inc. San Diego, CA
Schwartz	Richard	1991/ Ph.D.	Development and Design of Ex Vivo Bone Marrow Bioreactor Systems	Avron
Shen	Bing-Qian	1993/ POST	Optimizing Retroviral Titers	Metabolex, San Francisco, Research Scientist
Wilback (Smith)	Sharon	2004/ PhD	Applying kinetic constraints and regulatory logic to metabolic networks	Genomatica, Inc
Trakakis	Ionnannis	1986/ MS	Metabolic Dynamics in Cardiac Muscle	
Theile	Ines	Current		UCSD Doctoral Student
Varma	Amit	1993/ Ph.D.	Flux Balance Analysis of Escherichia Coli Metabolism	PDL BioPharma
Vo	Thuy	2006/Ph D.	H. pylori model	Amgen
Yeung	Matt	2007/P OST		Researcher, Imperial College
Zengler	Karsten	Current		UCSD Post Doc
Zylla	Brian	1994/ MS	Shear Sensitivities of Human Bone Marrow Cultures	Upjohn, Kalamazoo, MI, Product Development Engineer

## COURSES TAUGHT

COURSE	TOPIC	INSTITUTION
ChE 230	THERMODYNAMICS I.	Univ. Michigan
ChE 330	THERMODYNAMICS II.	Univ. Michigan
ChE 341	FLUID MECHANICS.	Univ. Michigan
ChE 342	HEAT AND MASS TRANSFER.	Univ. Michigan
ChE 417	BIOCHEMICAL TECHNOLOGY.	Univ. Michigan
ChE 460	CHEMICAL ENGINEERING LABORATORY II.	Univ. Michigan
ChE 466	PROCESS DYNAMICS AND CONTROL.	Univ. Michigan
ChE 516	DYNAMICS OF BIOLOGICAL SYSTEMS.	Univ. Michigan

ChE 542	GRADUATE TRANSPORT PHENOMENA.	Univ. Michigan
BioE 160A	METABOLIC ENGINEERING.	UCSD
BioE 160C	IN SILICO BIOOLOGY	UCSD
BioE 166A	TISSUE ENGINEERING	UCSD
BioE 241B	TISSUE ENGINEERING; METHODS.	UCSD
BioE 241C	TISSUE ENGINEERING; SPECIAL TOPICS.	UCSD
BioE122A	BIOSYSTEMS AND CONTROLS	UCSD
BioE 123	SYSTEMS BIOLOGY AND BIOENGINEERING	UCSD
BioE 203	BIOINFORMATICS III	UCSD
BioE 211	SYSTEMS BIOLOGY AND BIOENGINEERING I	UCSD
BioE 212	SYSTEMS BIOLOGY AND BIOENGINEERING II	UCSD
BioE 213	SYSTEMS BIOLOGY AND BIOENGINEERING III	UCSD

## PATENTS

1. U.S. Patent No. 5,399,493, Methods and compositions for the optimization of human hematopoietic progenitor cell cultures
2. U.S. Patent No. 5,437,994, Method for the ex vivo replication of stem cells, for the optimization of hematopoietic progenitor cell cultures, and for increasing the metabolism, GM-CSF secretion and/or IL-6 secretion of human stromal cells
3. U.S. Patent No. 5,459,069, Device for maintaining and growing human stem and/or hematopoietics cells
4. U.S. Patent No. 5,534,423, Methods of increasing rates of infection by directing motion of vectors
5. U.S. Patent No. 5,605,822, Methods, compositions and devices for growing human hematopoietic cells
6. U.S. Patent No. 5,614,378, Photobioreactors and closed ecological life support systems and artificial lungs containing the same
7. U.S. Patent No. 5,616,487, Stabilized retrovirus compositions
8. U.S. Patent No. 5,635,386, Methods for regulating the specific lineages of cells produced in a human hematopoietic cell culture
9. U.S. Patent No. 5,646,043, Methods for the ex vivo replication of human stem cells and/or expansion of human progenitor cells
10. U.S. Patent No. 5,654,185, Methods, compositions and apparatus for cell transfection
11. U.S. Patent No. 5,670,147, Compositions containing cultured mitotic human stem cells
12. U.S. Patent No. 5,670,351, Methods and compositions for the ex vivo replication of human hematopoietic stem cells
13. U.S. Patent No. 5,672,494, Methods of increasing rates of infection by directing motion of vectors
14. U.S. Patent No. 5,688,687, Bioreactor for mammalian cell growth and maintenance
15. U.S. Patent No. 5,763,266, Methods, compositions and devices for maintaining and growing human stem and/or hematopoietics cells
16. U.S. Patent No. 5,804,431, Method, compositions and apparatus for cell transfection
17. U.S. Patent No. 5,811,274, Methods, compositions and apparatus for cell transfection
18. U.S. Patent No. 5,866,400, Methods of increasing rates of infection by directing motion of vectors
19. U.S. Patent No. 5,874,266, Targeted system for removing tumor cells from cell populations
20. U.S. Patent No. 5,888,807, Devices for maintaining and growing human stem and/or hematopoietics cells
21. U.S. Patent No. 6,143,535, Targeted system for removing tumor cells from cell populations
22. U.S. Patent No. 6,326,198 B1, Methods and compositions for the ex vivo replication of stem cells, for the optimization of hematopoietic progenitor cell cultures, and for increasing the metabolism, GM-CSF secretion and/or IL-6 secretion of human stromal cells

23. U.S. Patent No. 6,514,722 B2, Method and apparatus for selectively targeting specific cells within a cell population
24. U.S. Patent No. 6,524,797 B1, Methods of identifying therapeutic compounds in a genetically defined setting
25. U.S. Patent No. 6,667,034, Methods for regulating the specific lineages of cells produced in a human hematopoietic cell culture, methods for assaying the effect of substances on lineage-specific cell production, and cell compositions produced by these cultures
26. U.S. Patent No. 6,753,161 Optoinjection methods
27. U.S. Patent No. 6,804,385 Method and device for selectively targeting cells within a three-dimensional specimen
28. U.S. Patent No. 7,127,379 B2, Method for the evolutionary design of biochemical reaction networks

#### TECHNOLOGY TRANSFER ACTIVITIES

- |                                     |              |  |
|-------------------------------------|--------------|--|
| 1. AASTROM BIOSCIENCES              | Founded 1988 | <a href="http://www.aastrom.com">www.aastrom.com</a>       |
| 2. ONCOSIS                          | Founded 1997 | <a href="http://www.oncosis.com">www.oncosis.com</a>       |
| 3. ICELAND GENOMICS CORP.           | Founded 1998 | <a href="http://www.uvs.is">www.uvs.is</a>                 |
| 4. GENOMATICA                       | Founded 1999 | <a href="http://www.genomatica.com">www.genomatica.com</a> |
| 5. CYNTELLECT<br>(Formerly Oncosis) | Founded 2001 | <a href="http://www.cyntellect.com">www.cyntellect.com</a> |

# Toward a Science of Metabolic Engineering

JAMES E. BAILEY

**Application of recombinant DNA methods to restructure metabolic networks can improve production of metabolite and protein products by altering pathway distributions and rates. Recruitment of heterologous proteins enables extension of existing pathways to obtain new chemical products, alter posttranslational protein processing, and degrade recalcitrant wastes. Although some of the experimental and mathematical tools required for rational metabolic engineering are available, complex cellular responses to genetic perturbations can complicate predictive design.**

THE METABOLIC ACTIVITIES OF LIVING CELLS ARE ACCOMPLISHED by a regulated, highly coupled network of ~1000 enzyme-catalyzed reactions and selective membrane transport systems. However, metabolic networks that evolved in natural settings are not genetically optimized for the objectives important in practical applications. Hence, performance of bioprocesses can be enhanced by genetic modification of the cells.

Metabolic engineering is the improvement of cellular activities by manipulation of enzymatic, transport, and regulatory functions of the cell with the use of recombinant DNA technology. The opportunity to introduce heterologous genes and regulatory elements distinguishes metabolic engineering from traditional genetic approaches to improve the strain. This capability enables construction of metabolic configurations with novel and often beneficial characteristics. Cell function can also be modified through precisely targeted alterations in normal cellular activities. Examples in the manipulation of protein processing pathways, as well as of pathways involving smaller metabolites, will be highlighted here.

At present, metabolic engineering is more a collection of examples than a codified science. Results to date promise future technological benefits, as well as contributions to basic science, agriculture, and medicine. However, many studies have shown the feasibility of metabolic engineering methods without achieving the yields, rates, or titers (final concentrations) required for a practical process. Most experiments explore changes in a single gene, operon, or gene cluster. After a new strain has been created by such a manipulation, limitations arise that can in principle be addressed by subsequent genetic manipulation. An iterative cycle of a genetic change, an analysis of the consequences, and a design of a further change, analogous to that articulated for protein engineering (1), can be used to find an optimized strain. The few cases to date in which such a metabolic engineering cycle has been implemented have achieved success. An emerging base of strategies, tools, and experiences will aid in identifying, implementing, and refining which particular set of

genetic manipulations is most effective in accomplishing a desired change in cellular function.

## Recruiting Heterologous Activities for Strain Improvement

Cloning and expression of heterologous genes can serve several useful purposes, including extending an existing pathway to obtain a new product, creating arrays of enzymatic activities that synthesize a novel structure, shifting metabolite flow toward a desired product, and accelerating a rate-determining step. Introduction of a functional heterologous enzyme or transport system into an organism can result in the appearance of new compounds that may subsequently undergo further reactions. Difficulties in anticipating these further reactions are a central limitation of metabolic engineering.

Expression of a heterologous protein does not guarantee appearance of the desired activity. The protein must avoid proteolysis, fold properly, accomplish any necessary assembly and prosthetic group acquisition, be suitably localized, have access to all required substrates, and not encounter an inhibitory environment. Despite these potential barriers to the successful recruitment of heterologous cellular activities, the number and scope of positive experiments encourage further application of this approach.

*Synthesis of new products is enabled by completion of partial pathways.* The genetic and metabolic diversity that exists in nature provides a collection of organisms with a spectrum of substrate assimilation and product synthesis capabilities. However, many natural strains are imperfect from an applied perspective. Their performance can sometimes be enhanced by extension of their native pathways. Native metabolites can be converted to preferred end products by the genetic installation of a few well-chosen heterologous activities (Table 1).

For example, the final precursor in a current commercial process for ascorbic acid (vitamin C) synthesis is 2-keto-L-gulonate (2-KLG). One route to 2-KLG involves two successive fermentations. The first converts glucose to 2,5-diketo-D-gluconate (2,5-DKG) in *Erwinia herbicola*; the second fermentation, carried out in a species of *Corynebacterium*, transforms 2,5-DKG to 2-KLG. Researchers devised a way to convert glucose to 2-KLG in a single fermentation step by cloning the *Corynebacterium* enzyme 2,5-DKG reductase, which catalyzes the 2,5-DKG to 2-KLG conversion, into *E. herbicola* (2). A similar goal was achieved for 7-aminocephalosporanic acid (7ACA), the precursor for several semisynthetic cephem antibiotics (3).

Posttranslational modifications can influence the function of proteins. The types of modifications that occur can be affected by expression of cloned protein processing enzymes. For example, expression in Chinese hamster ovary (CHO) cells of  $\beta$ -galactoside  $\alpha$ 2,6-sialyltransferase (4) allows the formation of sialyl  $\alpha$ 2,6-galactosyl linkages on its surface glycoproteins. These terminal glycosylation linkages are normally absent from proteins produced in this industrial

The author is the Chevron Professor of Chemical Engineering at the California Institute of Technology, Pasadena, CA 91125.

cell line, including cloned erythropoietin. Thus, this strategy should enable erythropoietin made in recombinant CHO cells to more closely resemble human erythropoietin, which is rich in these linkages (5). In another study, mouse cells displayed the human H blood group antigen after transfection with human DNA (6).

**Transferring multistep pathways: Hybrid metabolic networks.** The transfer of genes that encode entire biosynthetic pathways to a heterologous host can provide more industrially robust strains, enhance productivity, or permit the use of less costly raw materials. Moreover, such experiments are useful for exploring the regulation and function of a multistep metabolic pathway in a particular species.

Transferring entire antibiotic biosynthetic pathways to heterologous hosts has been facilitated by the clustering of the genes involved (7). Genes for the biosynthesis of actinorhodin were transferred from *Streptomyces coelicolor* to *Streptomyces lividans*, enabling the latter strain to produce actinorhodin (8). Subsequently, clustered erythromycin biosynthetic genes from *Streptomyces erythraeus* were transferred to *S. lividans*, which then synthesized an antibiotic indistinguishable from erythromycin A (9). *Escherichia coli* carrying this cloned gene cluster did not synthesize the antibiotic, possibly because of low transcriptional activity of *Streptomyces* promoters in *E. coli*. The fungi *Neurospora crassa* and *Aspergillus niger*, which normally do not produce  $\beta$ -lactam antibiotics, synthesized penicillin V after transformation with a cosmid containing *Penicillium chrysogenum* DNA that encoded enzymes in the penicillin biosynthetic pathway (10).

Polyhydroxybutyrate (PHB), a storage product sequestered in large amounts by some bacteria under growth-limiting, carbon source-excess conditions, is a biodegradable polyester that already has small-scale applications. *Alcaligenes eutrophus* can produce not only PHB but, when supplied with different precursors, can synthesize various polyhydroxyalkanoate copolymers as well (11). Metabolic engineering of the synthesis of these and related polymers should provide greater control over the nature and quantity of the polymer produced and should also offer alternative production organisms. The PHB synthesis operon from *A. eutrophus*, which encodes PHB polymerase, thiolase, and reductase activities, has been used to transform *E. coli* (12). As in *A. eutrophus*, this recombinant *E. coli* accumulates PHB when the nitrogen source is depleted; PHB concentrations in these cultures reach 50% of the dry cell weight.

Assembly of pathways for simultaneous degradation of chloro- and methylaromatics by combining and refining of cloned pathway segments and regulatory systems from several different organisms

exemplifies the iterative design of an effective hybrid organism (13). The ultimate strain thus far constructed contains five pathway segments obtained from three organisms. The biochemical and metabolic complexities of the degradation of mixed substrates and the resulting rationale behind each portion of this construction offer useful general perspectives on metabolic engineering strategies (14).

**Creating new products and new reactants.** Expression of biosynthetic genes for a secondary metabolite in a heterologous host that synthesizes its own different secondary metabolite can result in the construction of an array of enzymatic activities that yield novel products. Among the novel antibiotics that have been produced in recombinant strains of *Streptomyces* by such manipulations are mederhodins A and B and dihydrogratidin (15), 2-norerythromycins A, B, C, and D (16), and isovaleryl spiramycin (17).

Compounds new to the cell that result from a heterologous activity often undergo further reactions. In some cases, such as in the biosynthesis of indigo by *E. coli* that express *Pseudomonas putida* naphthalene dioxygenase (18), these subsequent reactions are essential components of the desired pathway. Another illustration of metabolic engineering to introduce a novel intermediate into a host involves recombinant *E. coli* that express the cloned tyrosinase gene from *Streptomyces antibioticus* (19). Synthesis by the recombinant *E. coli* strain of the pigment melanin, an ultraviolet light-absorbing compound with material and cosmetic applications, depends on a single critical catalytic step: the oxidation of tyrosine and L-dopa to dopaquinone by tyrosinase; the remaining reactions that yield melanin are apparently nonenzymatic. Melanin production is increased when another protein from *S. antibioticus* is coexpressed with tyrosinase. Although definitive evidence is not yet available, this second protein may provide a copper-donor function that activates apotyrasinase. Thus, increasing the expression of a cofactor-requiring protein as part of a metabolic engineering scheme may require the engineering of an increased supply of the cofactor as well.

Biodegradation of undesirable compounds can often be accomplished by host enzymes after a heterologous activity provides the initial attack on the target compound or compounds. For example, the expression of *Pseudomonas mendocina* toluene monooxygenase in *E. coli* enabled the efficient degradation of trichlorethylene, a suspected carcinogen and widespread pollutant (20). In *E. coli*, degradation can be induced by isopropyl-1-thio- $\beta$ -D-galactoside or by a temperature shift, rather than by toluene, as occurs in *P. mendocina*. In addition, the engineered *E. coli* has degradation kinetics (no competitive inhibition, as with toluene) and cosubstrate

**Table 1.** Heterologous activities recruited to alter small metabolite and protein end products. The original metabolite serves as the substrate for the synthesis of the new product through a pathway involving the new intermediate. It is difficult to prove that the inserted activity alone is responsible for

an altered phenotype. In each case discussed here, the observed change in cell function is consistent with the expected consequence of the newly installed gene or genes. *A. chrysogenum*, *Acetomonas chrysogenum*; GDP, guanosine diphosphate; *F. solani*, *Fusarium solani*; *P. diminuta*, *Pseudomonas diminuta*.

Host organism	Original metabolite	Heterologous enzymes added (source organism)	New product (new intermediate)
<i>E. herbicola</i> (2, 21, 22)	2,5-DKG	2,5-DKG reductase ( <i>Corynebacterium</i> )	2-KLG
<i>A. chrysogenum</i> (3)	Cephalosporin C	D-Amino acid oxidase ( <i>F. solani</i> ), cephalosporin acylase ( <i>P. diminuta</i> )	7ACA [7-(5-carboxy-5-oxopentanimido)-cephalosporanic acid]
CHO cells (4)	Terminal $\beta$ -galactosyl residues in N-acetylglucosamine sequences	$\beta$ -Galactoside $\alpha$ 2,6-sialyltransferase (rat)	Sialyl $\alpha$ 2,6-galactosyl linkages
Mouse L cells (6)	Unsubstituted type II, N-acetylglucosamine glycoconjugate end groups	GDP-L-fucose: $\beta$ -D-galactoside 2-O-L-fucosyltransferase (A431 human cell line)	H fucosyl $\alpha$ 1-2 galactosyl linkages



Systematic genetic manipulation of protein processing pathways has proven effective in increasing the quantity of active protein recovered. Overexpression of the *E. coli* chaperone proteins GroES and GroEL provided a five- to tenfold increase in assembled cyanobacterial Rubisco (D-ribulose-1,5-bisphosphate carboxylase/oxygenase) enzyme coexpressed in *E. coli*. In vitro studies of interactions among Rubisco and the GroE proteins implicate  $Mg^{2+}$ -ATP as a requirement for assembly (32). The failure to achieve assembly of Rubisco from higher plants in altered *E. coli* signals future challenges in the transfer of heterologous protein processing pathways (33). Other challenges include genetic manipulations of processing pathways of bacteria that alter the solubility of recombinant proteins (34). In addition, opportunities exist for extending such strategies to eukaryotic hosts (35).

**Transfer of promising natural motifs:** *Vitreoscilla hemoglobin*. Because of the constant drive toward maximum cell densities to maximize volumetric productivity, growth and product synthesis in many industrial processes are limited by oxygen supply. The Gram-negative aerobic bacterium *Vitreoscilla*, which lives in poorly aerated environments, synthesizes increased quantities of a hemoglobin molecule in oxygen-limited cultures (36). Although the function of this protein in its natural host has not been established, this pattern of regulation of expression, combined with the oxygen-binding and release characteristics of the protein, suggest a possible beneficial physiological activity in poorly oxygenated environments.

Motivated by this hypothesis and the premise that this beneficial function might be genetically transferred to industrial microorganisms, the gene for *Vitreoscilla hemoglobin* (VHB) was cloned and expressed in *E. coli* (37). *Escherichia coli* that carried a single copy of this gene integrated in the chromosome synthesized total cell protein more rapidly than an isogenic wild-type strain in oxygen-limited cultivations (Fig. 2), a response attributed to an increased efficiency of net ATP synthesis in the hemoglobin-expressing strain (38). Facilitation of oxygen transfer to the respiratory center (39) and modification of some aspect of cellular redox chemistry (38) have been suggested as contributing mechanisms for these phenomena. Coexpression of VHB increases the expression of cloned  $\beta$ -galactosidase, chloramphenicol acetyltransferase (CAT) (38), and  $\alpha$ -amylase (40) by 1.5- to 3.3-fold relative to controls in oxygen-limited *E. coli* cultures, probably as a result of enhanced net ATP synthesis.

Aeration of bioreactors used in the synthesis of antibiotics is

frequently complicated by the thick broths that result from growth of filamentous fungi and *Streptomyces*. Success of the strategy for enhancing aerobic metabolism in other bacteria prompted cloning and intracellular expression of VHB in two different *Streptomyces* species (41). *Streptomyces lividans* with a multicopy hemoglobin expression plasmid achieved final cell densities up to 54% greater than the untransformed host in shake-flask cultivations. The presence of cloned intracellular VHB in *S. coelicolor* markedly increased secondary metabolite accumulation, without affecting cell growth relative to a control strain containing a mutated VHB gene (Fig. 2).

These examples suggest a general genetic strategy for addressing stresses and corresponding productivity limitations encountered in bioprocessing: after identifying a response in nature to a similar stress (most likely involving a different organism), genes that specify that response can be transferred to the organism of choice.

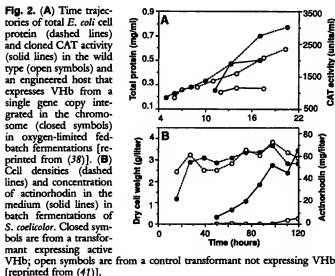
## Redirecting Metabolite Flow

Typically the route of reactions to a desired product passes several forks where intermediates can enter alternative pathways. At such bifurcations of metabolite flow, a common resource—for example, substrate, enzyme, transport system, or ribosome—contributes to two or more parallel processes. Maximizing product formation requires that the desired route at each fork be made a priority and that traffic in alternative pathways be minimized to the extent possible without decreasing cell viability.

**Directing traffic toward the desired branch.** Amplification of the activity initiating a desired process at a fork in a metabolic flow is a common strategy of metabolic engineering. Whereas isolation of mutant enzymes that are desensitized to feedback repression was achieved with classical methods, such mutants may now be obtained more rapidly with the use of cloned genes. This approach also avoids the complication of uncharacterized additional mutations that are often obtained with classical, whole-cell mutagenesis.

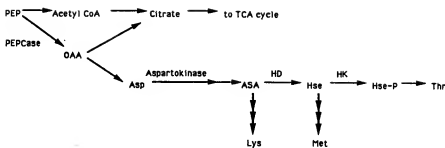
The past decade has seen a new generation of strain improvements in amino acid-producing coryneform bacteria with metabolic engineering (also called molecular breeding) (42, 43). Central to the success achieved was the development of new vectors and transformation procedures.

Genetic engineering of improved threonine production by *Brevibacterium lactofermentum* illustrates some of the strategies useful in redirecting metabolite flow to the desired product. Figure 3 presents an abbreviated diagram of the reactions involved in the synthesis of the aspartate family of amino acids and a few key reactions that feed into the synthesis pathway for this family. Homoserine dehydrogenase (HD) was amplified by cloning and transformation into a threonine- and lysine-producing mutant (designated M-15). This mutant organism was selected for its lack of feedback inhibition of aspartokinase by threonine and lysine and of HD by threonine (44). The respective final concentrations of threonine, homoserine, and lysine from benchtop fermentations were 25.0, 2.8, and 1.1 g/liter for the recombinant strain compared to 17.5, 0.5, and 12.1 for M-15. Subsequent further engineering to coexpress cloned homoserine kinase (HK) with HD further increased the final threonine concentration to 33 g/liter and reduced homoserine and lysine levels, relative to the strain with cloned HD alone (45). In another study, the coryneform gene for HD was mutagenized to eliminate feedback inhibition by threonine. Introduction of this mutated HD gene into a lysine producer shifted the final lysine concentration from 65 g/liter to 4 g/liter and the final threonine concentration from 0 g/liter to 52 g/liter (43). Threonine production by M-15 was increased 12% by the expression of cloned phosphoenolpyruvate (PEP) carboxylase (PEPCase) (46). This manipulation was motivated by





**Fig. 3.** Pathways of the biosynthesis of the aspartate family of amino acids. Metabolite abbreviations are as follows: acetyl CoA, acetyl coenzyme A; TCA, tricarboxylic acid cycle; Asp, aspartate; ASA, aspartate semialdehyde; Hse, homoserine; Lys, lysine; Met, methionine; Hse-P, O-phospho-homoserine; and Thr, threonine.



the researchers' desire to increase oxaloacetate (OAA) production and thereby to increase carbon flow into amino acid production. Further improvements in rates of amino acid synthesis and yields will depend on a better understanding of mechanisms of regulation of gene expression and metabolite flow in these bacteria (47, 48).

Because of metabolic engineering, *E. coli* has become an industrially important producer of amino acids. Transformation by multicopy plasmids that contain tryptophan (49) and threonine (50) biosynthetic genes have increased production of these amino acids. A project to engineer phenylalanine production in *E. coli* showed that overexpression of some genes in the phenylalanine biosynthetic pathway could cause a decrease in phenylalanine production and that inducible excision vector technology can be used to manipulate the biosynthesis of tyrosine, an inhibitor of the desired pathway (51). An intermediate in a metabolic pathway can be overproduced by combining a mutation that blocks that intermediate's use by the cell and by genetic augmentation of precursor flow into that pathway; although this concept has been used extensively in classical genetic production of organisms that are amino acid overproducers, it can be implemented in other contexts by metabolic engineering (52).

The gene *eryF* in *Saccharopolyspora erythrae* encodes the first enzyme in the pathway from 6-deoxyerythronolide B to the antibiotic erythromycin. After the targeted disruption of this gene using an integrative plasmid, 6-deoxyerythronolide B was converted to an erythromycin derivative that is more stable at the low pH of the stomach (53).

Because enzyme activities involved in secondary metabolite production are regulated at both the gene and protein levels, identifying genetic changes that accelerate synthesis of these metabolites is challenging. One successful strategy is based on measuring the biosynthetic pathway intermediate concentrations in the growth medium. Relatively high extracellular concentrations of the intermediate penicillin N suggested that the activity that converts this intermediate to cephalosporin C (encoded in *cefEF*) may limit the rate of the overall pathway (54). Thus, expression of *cefEF* was elevated through increased gene dosage in a production strain of *Cephalosporium acremonium*. This recombinant fungus exhibited a 15-fold reduction in penicillin N production and an increase of ~15% in cephalosporin C production.

Routing through protein processing pathways has also been altered by manipulation of host genes. The overproduction lethality commonly observed with exported  $\beta$ -galactosidase fusion proteins in *E. coli* is suppressed by the overproduction of *E. coli* *prfF* (55), and the expression of *E. coli* *DnaK* enables export of *lacZ*-hybrid proteins that are otherwise confined to the cytoplasm (56). An  $\text{NH}_2$ -terminal methionine often differentiates cloned polypeptides synthesized in *E. coli* from their native human counterparts. Coexpression of cloned *E. coli* methionine aminopeptidase with human interleukin-2 in *E. coli* has substantially reduced the fraction of product with methionine at its  $\text{NH}_2$ -terminus (57). Observation of large quantities of a variant of human tissue plasminogen activator

(tPA) associated with GRP78 in the rough endoplasmic reticulum of CHO cells suggested that GRP78 binding was a rate-limiting step in tPA secretion. (GRP78 is the 78-kD glucose-regulated protein, one of the stress-response proteins.) Coexpression of antisense GRP78 message resulted in smaller quantities of GRP78 and faster tPA secretion (58).

A quantitative study was conducted of *S. cerevisiae* isolates that contain different numbers of the phosphoglycerate kinase (PGK) gene (*PGK1*). In some cases a 10 to 15% increase in PGK activity gave rise to a higher (30%) overall cell mass yield when the yeast were grown on glucose. However, in another construct that contained more copies of *PGK1*, yield was depressed by 40% (59). These results show the importance of fine-tuning the amount of gene or enzyme amplification to achieve the desired benefit.

**Reducing competition for a limiting resource.** Computer simulations with a detailed single-cell model suggested that competition between vector- and host-encoded messages for a common pool of ribosomes could limit cloned gene expression (60), a prediction consistent with experimental observations of reduced ribosome content in recombinant *E. coli* that contain more plasmids per cell (61, 62). By expressing a cloned mutant 16S ribosomal RNA, a population of ribosomes is created that is specialized for expression of only those cloned gene transcripts that bear a corresponding mutation in the Shine-Dalgarno sequence (63). Large amounts of messenger RNA transcribed from the cloned gene will not interact with the primary population of native ribosomes; therefore the expression of cloned genes will not interfere with the simultaneous expression of host cell genes important for protein synthesis. With this approach, expression of cloned  $\beta$ -galactosidase was increased by 35%. A 30% reduction in specific growth rate occurred after expression of the specialized 16S RNA; however, no growth rate reduction was observed on induction of cloned  $\beta$ -galactosidase synthesis (64).

**Revising metabolic regulation.** Positive control genes have been found so far in the biosynthetic gene clusters for actinorhodin, bialaphos, streptomycin, and undecylprodigiosin, which are all secondary metabolites produced by *Streptomyces* species. Cloning additional copies of an activator gene in the wild-type host can substantially increase antibiotic production, as indicated for undecylprodigiosin (8).

Modification of the regulation of expression of maltose permease and maltase is the basis for a metabolically engineered baker's yeast intended to reduce the time for leavening of sweet doughs (65). Glucose normally represses expression of these proteins and thereby blocks simultaneous maltose utilization. The engineered strain uses constitutive yeast promoters for these enzymes to enable simultaneous uptake and catabolism of both sugars. This is one of a few examples in which a transport system has been manipulated.

A genetic engineering strategy to stimulate  $\text{CO}_2$  production by bakers' yeast seeks to consume ATP (66). This could relieve ATP inhibition of phosphofructokinase (PFK) and pyruvate kinase, two regulatory enzymes in sugar catabolism. A futile cycle with PFK

was created by expressing cloned yeast fructose 1,6-bisphosphatase (FBPase) from a yeast glycerophosphate dehydrogenase promoter that is induced by glucose; FBPase is not normally expressed in the presence of high glucose concentrations. This yeast strain produced 20 to 25% more CO<sub>2</sub> than the wild type.

In order to construct a pathway for processing the pollutant 4-ethylbenzoate, it was necessary to alter the regulation of the alkylbenzoate degradation pathway encoded on the *Pseudomonas* TOL plasmid (67). Originally 4-ethylbenzoate did not induce transcription of the crucial *meta* operon. Mutations were introduced into the positive regulator of *Pm* (the promoter of the *meta* operon) that enabled *Pm* activation by 4-ethylbenzoate.

## Completing the Metabolic Engineering Cycle: Potentials and Perils of Rational Design

The iterative cycle of genetic modification, analysis of the metabolic consequences of this change, and choice of the next genetic modification has been successfully implemented in a few instances with promising results. Contemporary concepts and technologies for each function in this cycle are summarized next.

**Cloning in industrial strains.** The lack of suitable vectors and methods for the introduction of exogenous DNA limits the application of metabolic engineering in many important industrial organisms (68). Electroporation and conjugation have proven useful in introducing DNA into diverse organisms. The stable propagation of cloned genes remains problematic even in such a well-studied system as *Bacillus subtilis* and is apparently a result of the error-prone rolling-circle replication mechanism used by many plasmids in Gram-positive bacteria. Extensive rearrangements and deletions of both chromosomal and plasmid DNA occur frequently in some species, complicating their systematic manipulation. Restriction (cleavage) of heterologous DNA is a limitation in efficient engineering of many cells of practical interest.

The hurdles to be surmounted in developing the necessary genetic tools are illustrated in research that is establishing a foundation for engineering the complex catabolic metabolism of *Clostridium acetobutylicum*. This bacterium is the basis for the biological production of the industrial chemicals acetone and butanol. Efficient transformation of this organism required optimization of an electroporation protocol, and it was discovered that, because of a clostridial restriction enzyme system, *E. coli* is not a suitable organism for the cloning of clostridial DNAs, whereas *B. subtilis* is (69). Technology for chromosomal integration should soon follow, as several *C. acetobutylicum* genes have now been cloned.

**Dissecting physiological responses.** For the most effective design of a subsequent genetic manipulation, it is useful to know the concentrations of intracellular proteins and metabolites. The concentrations of many cellular proteins can be determined in principle from two-dimensional gel electrophoresis, but data bases are necessary to identify individual proteins (70). In vitro assays of changes in activities of key enzymes have been widely applied.

A broad spectrum of analytical methods can be applied for determining metabolite concentrations. The measurement of concentrations and in some cases of fluxes in particular pathways of interest can often be aided by the application of isotopically labeled precursors. For example, with the use of labeled acetate and glutamate, along with quasi-steady state conservation equations for intracellular metabolites, the velocities for carbon flow through *E. coli* growing on acetate can be determined (71).

Nuclear magnetic resonance (NMR) spectroscopy has been applied to estimate metabolite concentrations in whole cells, cell extracts, and growth media (72). For example, <sup>31</sup>P NMR measure-

ments of *S. cerevisiae* cells converting glucose to end products under anaerobic conditions, in concert with a methodology for extracting individual component information from the sugar phosphate portion of the spectrum, provided estimates of metabolite concentrations that were essential for analysis of the pathway (73). The time and instrumentation required to evaluate metabolite concentrations presently limits rational metabolic engineering.

**Design principles and cell models: Coping with complexity and coupling.** No universal principles have emerged from metabolic engineering research to guide the choice of the next useful genetic alteration. Attempts to address these problems with artificial intelligence have shown that there is no substitute for knowledge of the pathways involved, their regulation, and their kinetics. Some useful approaches include measurements of intermediate concentrations to indicate possible rate-determining reactions, genetic transfer of natural stress response motifs, and applications of organisms that can be used over wide ranges of temperature and pH (74).

Alternatively, if a mathematical description of the system is available, sensitivity analysis can be applied to calculate the expected response of the pathway to changes in the individual steps or pathway segments. An advantage of such an approach is its simultaneous determination of the sensitivities of the desired flux to many different participating reactions, permitting the identification of situations in which several genetic modifications in concert are required to achieve a desired response.

A body of theoretical developments known as metabolic control theory is well suited to the requirements of rational metabolic engineering (75). A central result provided by this theory is a sensitivity calculation that provides the flux control coefficients, defined as the fractional changes of flux expected for a unit fractional change in the amount of each enzyme participating in a given pathway. In addition, it is possible to evaluate the sensitivity of flux through the pathway to individual parameters in kinetic expressions for each of the enzymes, thereby providing guidance for useful protein engineering to accelerate the pathway. Analysis of several simple examples that involve unbranched sequences of reactions showed that sole control of flux by any single step (in other words, the existence of a single, rate-limiting step) is in general not expected. Instead, the flux through the pathway is usually influenced by the activities of several individual steps. This result, augmented by specific model calculations outlined below, provides motivation for sequential improvement of metabolic pathways.

In one of the few cases in which detailed kinetic expressions for each step in the reaction network, as well as the concentrations of all substrates and effectors, are known or estimated, flux control coefficients were determined for nongrowing yeast converting glucose to ethanol and other end products (76). Several general points are suggested by this investigation. First, the sensitivity of pathway flux to individual step changes depends on the environment in which the cell is grown. It is therefore important to carry out modeling and measurement under the expected industrial conditions. Second, flux control can be extraordinarily sensitive to some parameters such as intracellular pH. Third, in any system with interacting pathways (and it is difficult to envision any case where this does not occur), the most general version of metabolic control theory must be used (77). Such coupling is extensive in the usual case of growing cells, where the pathway of interest interacts with all of the other metabolic processes in the cell. A strategy for accomplishing flux control coefficient calculations in this situation has been presented (78). Finally, calculations with the kinetic model formulated in a yeast biocatalysis study indicate that amplification of the activity of one enzyme results in a shift of flux control to other steps in the pathway. New theory that presumes linear approximations for all rate expressions provides estimates of flux control coefficients,

without requiring knowledge of kinetics and using time-resolved metabolite concentration measurements instead (79); the practical merits of this approach have not yet been evaluated.

Both sensitivity to small changes and simulations of responses to large changes in intracellular activities can be calculated from a detailed and reliable mathematical model of the cell. Large quantities of biological information have been integrated into computer models for single cells (60, 80) and several molecular control systems (81). These have successfully simulated consequences of several genetic and environmental changes. Although useful initial directions for genetic improvement have been suggested by such models, they have not yet been used as the central tool in an iterative metabolic engineering study. In spite of their obvious limitations, these mathematical structures are the only way that the net consequences of simultaneous, coupled, and often counteracting processes can be simulated and evaluated consistently and quantitatively.

**Minimizing response cascades.** Unanticipated cell responses to a genetic modification may complicate rational practice of the metabolic engineering cycle. Introduction of a cloning vector alone may result in a large cascade of metabolic changes, many of which are difficult to anticipate. For example, the introduction of multicopy plasmids into *E. coli*, even without overexpression of a cloned product, has been shown to cause substantial changes in growth rates, cell cycle regulation, amounts of many individual proteins, glucose uptake, and carbon catabolite production rates (82). Transformation of yeast with multicopy plasmids can introduce lesions that persist after the plasmids have been eliminated from the transformants (83). Different mammalian cell clones transfected by the same vector often exhibit different growth rates and cell sizes. Therefore, introduction of a desired genetic change should be carefully configured to minimize perturbation of the host, using the lowest gene dosage and lowest expression level that give the desired result. The apparatus used for selecting the modified strain should also be carefully considered. For example, ampicillin resistance used for the maintenance of many laboratory recombinant *E. coli* strains is provided by cloned  $\beta$ -lactamase. This precursor must be processed at the cytoplasmic membrane in competition with host cell preprotein, which often results in a major physiological disruption.

Even if the genetic manipulation is accomplished in a relatively well-controlled fashion, the regulatory apparatus of the cell at both the gene and protein levels may confound the intended change or even alter cellular activities. For example, amplification of citrate synthase in *E. coli* did not increase the flux through the citric acid cycle because of a compensating modulation of the activity of isocitrate dehydrogenase (84). Expression of even low concentrations of unnatural proteins can activate stress responses, influencing many cell functions (85). Anticipating and accounting for such regulatory responses to genetic intrusions are fundamental challenges for the future.

#### REFERENCES AND NOTES

1. K. M. Ulmer, *Science* **219**, 666 (1983).
2. S. Anderson et al., *ibid.* **230**, 144 (1985).
3. T. Inoué et al., *Biotechnology* **9**, 188 (1991).
4. E. U. Lee, J. Roth, J. C. Paulson, *J. Biol. Chem.* **264**, 13848 (1989).
5. M. Takeuchi et al., *ibid.* **263**, 3657 (1988).
6. L. K. Ernst et al., *ibid.* **264**, 3436 (1989).
7. F. Malpartida, J. Niemi, R. Navarrete, D. A. Hopwood, *Gene (Amsterdam)* **93**, 91 (1990).
8. F. Malpartida and D. A. Hopwood, *Nature* **309**, 462 (1984).
9. R. Starzak, P. Matsushima, R. H. Balz, R. N. Rao, *Biotechnology* **4**, 229 (1986).
10. D. J. Smith, M. K. R. Burnham, J. Edwards, A. J. Earl, G. Turner, *ibid.* **8**, 39 (1990).
11. O. P. Peoples and A. J. Sinkey, in *Novel Biodegradable Microbial Polymers*, E. A. Davies, Ed. (Kluwer Academic, Dordrecht, Netherlands, 1990), pp. 191-202.
12. O. P. Peoples and A. J. Sinkey, *J. Biol. Chem.* **264**, 15298 (1989).
13. F. Rojo, D. H. Pieper, K.-H. Engesser, H.-J. Knackmuss, K. N. Timmis, *Science* **238**, 1395 (1987).
14. K. N. Timmis, F. Rojo, J. L. Ramos, in *Biotechnology and Biodegradation*, D. Karmali, A. Chakrabarty, G. S. Omen, Eds. (Portfolio, The Woodlands, Texas, 1990), pp. 61-80.
15. D. A. Hopwood et al., *Nature* **314**, 642 (1985); S. Omura, H. Ikeda, F. Malpartida, H. M. Kier, D. A. Hopwood, *Antimicrob. Agents Chemother.* **29**, 13 (1986).
16. J. B. McAlpine et al., *J. Antibiot.* **40**, 1115 (1987).
17. J. K. Ego, M. L. B. Huber, J. R. Turner, T. Goodson, B. E. Schoner, *Gene (Amsterdam)* **85**, 293 (1989).
18. B. D. Enslay et al., *Science* **222**, 167 (1983).
19. G. della-Cioppa, S. J. Garger, G. G. Sverlow, T. H. Turpen, L. K. Grill, *Biotechnology* **8**, 634 (1990).
20. R. B. Winters, K. M. Yen, B. D. Enslay, *ibid.* **7**, 282 (1989).
21. R. A. Lazarus et al., in *Biocatalysis*, D. Abramowitz, Ed. (Van Nostrand Reinhold, New York, 1990), pp. 136-155.
22. R. A. Lazarus et al., in *Proceedings of the Sixth International Symposium on Genetics of Industrial Microorganisms*, H. Heslot et al., Eds. (Société Française de Microbiologie, Strasbourg, 1990), vol. 2, pp. 1073-1082.
23. J. D. Windass et al., *Nature* **287**, 396 (1980).
24. L. O. Ingram and T. Conway, *Appl. Environ. Microbiol.* **54**, 397 (1988).
25. E. Bräu and H. Sahm, *Arch. Microbiol.* **144**, 296 (1986).
26. F. Altermann and L. O. Ingram, *ibid.* **55**, 1943 (1989).
27. J. S. Toland and R. F. Kim, *Appl. Environ. Microbiol.* **53**, 2039 (1987).
28. M.-L. Suilko, K. Blomqvist, M. Penttilä, R. Gisler, J. Knowles, *J. Biotechnol.* **14**, 285 (1990).
29. A. K. Koch, J. Reiser, O. Kappell, A. Fiechter, *Biotechnology* **6**, 1335 (1988).
30. J. D. Hamill et al., *Plant Mol. Biol.* **15**, 27 (1990).
31. T. Hashimoto, D. J. Yun, Y. Yamada, personal communication.
32. A. A. Gatenby, P. V. Vitazek, G. H. Lorimer, *Trends Biotechnol.* **8**, 354 (1990).
33. J. E. Rothman, *Cell* **59**, 591 (1989).
34. C. H. Schein, *Biotechnology* **7**, 1141 (1989).
35. J. L. Wise et al., *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9378 (1990).
36. S. J. Boorman and D. A. Webster, *J. Gen. Appl. Microbiol.* **28**, 35 (1982).
37. C. Khosla and J. E. Bailey, *Nature* **331**, 633 (1988); *Mol. Gen. Genet.* **214**, 158 (1988).
38. C. Khosla, J. Curtis, J. DeMolena, U. Rinas, J. E. Bailey, *Biotechnology* **8**, 849 (1990).
39. C. Khosla and J. E. Bailey, *J. Mol. Biol.* **171**, 5995 (1989).
40. M. Khorasani, D. A. Webster, B. C. Stark, *Plasmid* **24**, 190 (1990).
41. S. K. Magnolo et al., *Biotechnology* **9**, 473 (1991).
42. K. Sano, in *Proceedings of the Fourth European Congress on Biotechnology*, O. M. Neijssel, B. van der Meer, K. Ch. A. M. Luyben, Eds. (Elsevier Science, Amsterdam, 1987), vol. 4, pp. 735-747.
43. R. Katsumata et al., *ibid.*, pp. 767-776.
44. S. Nakamura et al., *Agri. Biol. Chem.* **51**, 87 (1987).
45. Y. Morinaga et al., *ibid.*, p. 93.
46. K. Sano, K. Ito, K. Sano, S. Nakamura, *ibid.*, p. 597.
47. M. T. Plettner, O. P. Peoples, J. A. C. Archer, A. J. Sinkey, in *Proceedings of the Sixth International Symposium on Genetics of Industrial Microorganisms*, H. Heslot et al., Eds. (Société Française de Microbiologie, Strasbourg, 1990), vol. 1, pp. 315-325.
48. G. N. Stephanopoulos, *Science* **252**, 1675 (1991).
49. S. Aiba, T. Imataka, H. Tsunokawa, *Biotechnol. Lett.* **2**, 525 (1980).
50. K. Miwa et al., *Agri. Biol. Chem.* **47**, 2329 (1983).
51. K. Backman et al., *Ann. N.Y. Acad. Sci.* **589**, 16 (1990).
52. K. M. Draths and J. W. Frost, *J. Am. Chem. Soc.* **112**, 1657 (1990).
53. J. M. Weber, J. O. Leung, S. J. Swanson, K. B. Idler, J. B. McAlpine, *Science* **252**, 114 (1991).
54. P. L. Skarud et al., *Biotechnology* **7**, 477 (1989).
55. D. R. Kline, G. J. Phillips, T. J. Silvary, *J. Bacteriol.* **172**, 185 (1990).
56. G. J. Phillips and T. J. Silvary, *Nature* **344**, 882 (1990).
57. A. Ben-Bassat et al., *J. Bacteriol.* **169**, 751 (1987).
58. J. Dorn, M. G. Kruse, R. J. Kaufman, *Mol. Cell. Biol.* **8**, 4063 (1988).
59. P. C. van der Aar et al., *Appl. Microbiol. Biotechnol.* **32**, 577 (1990).
60. S. W. Peretti and J. E. Bailey, *Biotechnol. Bioeng.* **29**, 316 (1987).
61. T. K. Wood and S. W. Peretti, *ibid.* **36**, 865 (1990).
62. S. Birnbaum and J. E. Bailey, *ibid.* **37**, 736 (1991).
63. A. Hui and H. A. de Boer, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4762 (1987).
64. T. K. Wood and S. W. Peretti, *Biotechnol. Bioeng.*, in press.
65. J. Hodgson, *Biotechnology* **8**, 281 (1990).
66. D. T. Rogers and E. Hiller, in *Proceedings of the Sixth International Symposium on Genetics of Industrial Microorganisms*, H. Heslot et al., Eds. (Société Française de Microbiologie, Strasbourg, 1990), vol. 1, pp. 533-544.
67. J. L. Ramos et al., *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8467 (1986).
68. H. Heslot et al., Eds., *Proceedings of the Sixth International Symposium on Genetics of Industrial Microorganisms*, H. Heslot et al., Eds. (Société Française de Microbiologie, Strasbourg, 1990); A. L. Demain and N. A. Solomon, Eds., *Biology of Industrial Microorganisms* (Benjamin-Cummings, Menlo Park, CA, 1985); C. L. Hershberger, S. W. Quencer, G. Hegeman, Eds., *Genetics and Molecular Biology of Industrial Microorganisms* (American Society for Microbiology, Washington, DC, 1989).
69. W. Cary, D. J. Petersen, G. N. Bennett, E. T. Papoutsakis, *Ann. N.Y. Acad. Sci.* **589**, 67 (1990).
70. T. A. Phillips, V. Vaughn, P. L. Bloch, F. C. Neidhardt, in *Escherichia coli and Salmonella typhimurium*, F. C. Neidhardt et al., Eds. (American Society for Microbiology, Washington, DC, 1987), pp. 919-966.

71. K. Walsh and D. E. Koshland, Jr., *J. Biol. Chem.* 259, 9646 (1984).
72. D. G. Gadian, *Nuclear Magnetic Resonance and Its Applications to Living Systems* (Oxford Univ. Press, Oxford, 1982); E. J. Fernandez and D. S. Clark, *Enzyme Microb. Technol.* 9, 256 (1987).
73. J. V. Shanks and J. E. Bailey, *Biotechnol. Bioeng.* 32, 1138 (1988).
74. T. D. Brock, Ed., *Thermophiles: General, Molecular and Applied Microbiology* (Wiley-Interscience, New York, 1986).
75. H. Kacer and J. A. Burns, *Symp. Soc. Exp. Biol.* 27, 65 (1973); R. Heinrich, S. M. Rapoport, T. A. Rapoport, *Prog. Biophys. Mol. Biol.* 32, 1 (1977).
76. J. L. Galazzo and J. E. Bailey, *Enzyme Microb. Technol.* 12, 162 (1990); *ibid.* 13, 363 (1991).
77. C. Rader, *J. Theor. Biol.* 135, 175 (1988).
78. P. M. Schlosser and J. E. Bailey, *Math. Biosci.* 100, 87 (1990).
79. J. P. Delgado and J. Liao, *Biotechnol. Prog.* 7, 15 (1991).
80. M. M. Domach et al., *ibid.* 26, 203 (1984).
81. S. B. Lee and J. E. Bailey, *Plasmid* 11, 151 (1984); *ibid.*, p. 166, *Biotechnol. Bioeng.* 26, 66 (1984); *ibid.*, p. 1372; *ibid.*, p. 1283.
82. J. E. Bailey, N. Da Silva, S. M. Peretti, J.-H. Seo, F. Srienc, *Ann. N.Y. Acad. Sci.* 469, 194 (1986).
83. N. Danhash, D. C. J. Gardner, S. G. Oliver, *Biotechnology* 9, 179 (1991).
84. K. Walsh, M. Schena, A. J. Flint, D. E. Koshland, Jr., *Biochem. Soc. Symp.* 54, 183 (1987).
85. M. Kosinski, U. Rinas, J. E. Bailey, unpublished data.
86. Supported by NSF and by the Catalysis and Biocatalysis Program of the Advanced Industrial Concepts Division of the U.S. Department of Energy. Advice and assistance in the preparation of this paper were provided by D. S. Clark, B. D. Ensley, T. Hashimoto, R. J. Kaufman, C. Khosla, R. A. Lazarus, A. J. Sinskey, K. N. Timmis, and colleagues at the California Institute of Technology.

# Network Rigidity and Metabolic Engineering in Metabolite Overproduction

GREGORY STEPHANPOULOS AND JOSEPH J. VALLINO\*

In order to enhance the yield and productivity of metabolite production, researchers have focused almost exclusively on enzyme amplification or other modifications of the product pathway. However, overproduction of many metabolites requires significant redirection of flux distributions in the primary metabolism, which may not readily occur following product deregulation because metabolic pathways have evolved to exhibit control architectures that resist flux alterations at branch points. This problem can be addressed through the use of some general concepts of metabolic rigidity, which include a means for identifying and removing rigid branch points within an experimental framework.

ALL ORGANISMS USE PRIMARY METABOLIC PATHWAYS to supply precursor metabolites and energy to anabolic pathways that synthesize cellular constituents that are necessary for growth and maintenance. In many industrial strains of microorganisms (as well as tissue and plant cultures), these anabolic pathways have been exploited for the overproduction of compounds (such as amino and nucleic acids, antibiotics, vitamins, enzymes, and proteins) that cannot be synthetically produced or for which it is not economical to do so. In general, a particular metabolite is overproduced by deregulating the pathway directly associated with the synthesis of that metabolite, or, more recently, by transforming a robust host organism (typically *Escherichia coli*) with the genes that encode for the synthesis of the desired product (1, 2). This approach, however, does not necessarily result in high product yields (defined as the moles of product formed per mole of substrate consumed) since carbon flux distributions at key branch points (nodes) in the primary metabolism [such as glycolysis, tricarboxylic

acid (TCA) cycle, and pentose phosphate pathway] must often be radically redirected from the flux distributions that are normally associated with balanced growth. Such metabolic flux alterations are often directly opposed by mechanisms for controlling enzyme activity that have evolved to maintain flux distributions that are optimal for growth. We refer to this inherent resistance to flux alterations as metabolic or network rigidity and to the genetic modifications of specific nodes in the primary metabolism for the purpose of enhancing yield and productivity as metabolic engineering (3). Although genetic manipulations can now be readily performed, there are relatively few accounts of successful metabolic flux alterations because of the complex, nonlinear nature of the metabolic control architectures.

The nature and types of metabolic rigidity are reviewed in this article along with methods to identify and possibly circumvent such undesirable nodal controls. The overproduction of lysine by *Corynebacterium glutamicum* [and related strains (4)] is used as a vehicle to illustrate key points because of: (i) the lack of compartmentalization in bacteria; (ii) the need for significant flux alterations to optimize lysine biosynthesis; and (iii) the apparent marginal success of mutation-selection (5, 6) or genetic engineering (7) techniques used to that end. The concepts, however, are of general value, and the methods are applicable to other metabolic products as well.

## Basis of Metabolic Rigidity

Although intracellular metabolite concentrations can fluctuate during growth, on average, the distributions of the major cellular groups (proteins, RNA, DNA, lipids, and so forth) remain relatively proportional to one another throughout balanced growth (8). In fact, metabolites and energy required to synthesize an *E. coli* cell have been calculated on the basis of its known composition (9). In order to preserve this regularity in cellular composition, the primary metabolism has evolved coordination of pathway control, such that building-block metabolites, energy [such as adenosine triphosphate (ATP)], and biosynthetic reducing power [such as nicotinamide adenine dinucleotide phosphate (NADPH)] are synthesized in approximate stoichiometric ratios during balanced growth. Al-

The authors are in the Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

\*Present address: Scripps Institution of Oceanography, University of California at San Diego, La Jolla, CA 92093-0202.

## IN MEMORIAM

### Jay Bailey (1944–2001)

Jay Bailey passed away in Zurich on 9 May 2001 at the age of 57. Jay received his education at Rice University, graduating with a B.A. in 1966 and Ph.D. in 1969, both in chemical engineering. After a short period with Shell Development he joined the Chemical Engineering Faculties of the University of Houston in 1971 and Caltech in 1980. In 1992 Jay was appointed Professor of Biotechnology at the Swiss Federal Institute of Technology (ETH) in Zurich.

In his early years Jay studied extensively the dynamics of chemical reactions and reaction networks with particular emphasis on the origin and interpretation of autonomous and forced oscillations of chemically reacting systems. During the 1970s, his interest shifted gradually to biological systems and this shift was culminated with the publication in 1977 of the landmark textbook *Biochemical Engineering Fundamentals* (coauthored with D. F. Ollis). At the time of his passing, Jay had coauthored approximately 400 publications, mostly in the field of biotechnology, including many seminal papers and visionary commentaries in leading journals.

Besides his outstanding contributions to biotechnology, Jay was instrumental in defining and building the foundations of metabolic engineering. He is arguably the first engineer who, in the early 1980s, embraced and promoted genetic engineering as an enabling new technology for improving cellular biocatalysts for industrial processes. This naturally led to the need to study the behavior of bioreaction networks in their entirety, analyze metabolic flux and flux control, and rigorously describe the physiology of wild-type and recombinant microorganisms. Although some of these questions had been investigated before by biochemists and chemical engineers, their sum total emerged as a distinct new field, metabolic engineering. Jay played a key role in setting the foundations and expanding this field.

Until the last day before being admitted to the hospital, Jay was actively working on the program of the 4th Conference of Metabolic Engineering, which he would have chaired in October of 2002. His plans for the conference program as well as his more recent scientific writings reveal a broader vision for metabolic engineering: First, metabolic engineering is at the forefront of functional genomics due to its efforts to describe the cellular physiology of wild-type microorganisms and recombinants with well-defined genetic backgrounds. Second, metabolic engineering pro-

vides an ideal *integrating platform* of genomic and physiological information and data by taking a holistic view of metabolic networks and cellular physiology for the identification of targets for genetic manipulation. As such, it is contributing concepts and methods of importance to systems biology. Finally, Jay expanded metabolic engineering beyond the original context of industrial strain improvement to also include mammalian cells, tissues, and medical applications.

Jay leaves a very rich legacy to the biotechnology and biochemical engineering communities. He mentored a school of distinguished students with brilliant careers in industry and academia and, through his teaching and writings, ushered biochemical engineering into the modern era of cellular and molecular biotechnology. Most importantly, he upheld the highest standards in education and research. These contributions make Jay undoubtedly the most influential biochemical engineer of modern time.

Jay was internationally recognized for his work in diverse areas of chemical engineering and biotechnology. He won numerous honors and awards in his career. The most recent was the First Merck Award in Metabolic Engineering. His award acceptance speech summarized his vision for metabolic engineering and formed the basis of the first perspective article in this journal (*Metab. Eng.* 3, 111–114, 2001). We are pleased to be able to reprint below the rendition of Bob Dylan's song *They Are A Changin'* by singer-poet Jay Bailey.

In Jay's honor, *Metabolic Engineering* is instituting a Young Investigator Best-Paper Award in Metabolic Engineering to be awarded every 2 years to the author(s) of an outstanding paper published in *Metabolic Engineering*. Details about the administration of the award will be provided in a future issue.

*The (Metabolic Engineering) Times,  
They Are A 'Changin'*

(Dylan/Bailey)

*Come gather 'round people wherever you roam,  
And admit that the waters around you have grown,  
And accept it that soon you'll be drenched to the bone.*

*If your time to you is worth saving,  
Then you'd better start swimming or you'll sink like a stone,  
For the times, they are a changin'.*

(Bob Dylan)



*Come gather Metabolic Engineers 'cross the land  
At ME III we'll take command  
Of cells that are too slow to produce or grow.  
If it's higher fluxes you're needin'  
Then we'll shift the controls, and block bad outflows.  
For the times, they are a changin'.*

*Do you need a new molecule or neutraceutical  
The Metabolic Engineer has the answers for you.  
We'll import new pathways, and shuffle them too.  
Is your lead compound library fadin'?  
We'll give new adducts to your old natural products  
For the times, they are a changin'.*

*Rational or random, which way is best?  
Solving the problem passes the test.  
Complex responses confuse the quest.  
More genetic and array technologies  
Will give us insights to networks' delights.  
For the times, they are a changin'.*

*Genomes are in hand, the sequences there,  
An amazing resource that we all share.  
Genes and controllers, bioinformatics tells us where.  
But how is all of this workin'?  
Let's decipher a yeast, understand that at least.  
For the times, they are a changin'.*

*How is phenotype controlled by the genes?  
Nobody knows, least of all the machines.  
Medicine will thrive if we can discover the means,  
To merge our knowledge and information  
And find genes' intent and control by environment.  
For the times, they are a changin'.*

*Metabolic Engineers have all the tools—  
Biology, computing, and engineering rules,  
Knowledge, experience, perspective on detail.  
Let's help Metabolic Genomics to set sail.  
Opportunity's here . . . but now it's time for a beer!  
For the times, they are a changin'.*

(Jay Bailey)  
October 2000

Videocassettes of Jay's Merck Award talk at the 3rd Metabolic Engineering Conference are available through the Engineering Foundation. Contact Ms. Barbara Hickernel at 212-591-7836 or [engfnd@aol.com](mailto:engfnd@aol.com).

Gregory Stephanopoulos  
Co-Editor, *Metabolic Engineering*

# Engineering of Molecular and Cellular Biocatalysts: Selected Contributions by James E. Bailey

Jonathan S. Dordick,<sup>1</sup> Chaitan Khosla<sup>2</sup>

<sup>1</sup>Department of Chemical Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180-3590; telephone: 518-355-4062; fax 518-276-2207; e-mail: dordick@rpi.edu

<sup>2</sup>Departments of Chemical Engineering and Chemistry, Stanford University, Stanford, California

DOI: 10.1002/bit.10040

**Abstract:** James (Jay) E. Bailey was a pioneer in biotechnology and biochemical engineering. During his 30 years in academia he made seminal contributions to many fields of chemical engineering science, including catalysis and reaction engineering, bioprocess engineering, mathematical modeling of cellular processes, recombinant DNA technology, enzyme engineering, and metabolic engineering. This article celebrates some of his contributions to the engineering of molecular and cellular biocatalysts, and identifies the influence he had on current and future research in biotechnology. © 2002 Wiley Periodicals, Inc. *Biotechnol Bioeng* 79: 490–495, 2002

**Keywords:** James Bailey; contributions to biotechnology and biochemical engineering; biocatalysis; metabolic engineering

## INTRODUCTION

James (Jay) E. Bailey, a pioneer in biotechnology and biochemical engineering, succumbed to metastatic cancer on May 9, 2001. After receiving his undergraduate and doctoral degrees at Rice University and a brief stint at Shell Development (Houston), he started his academic career in 1971 at the University of Houston. In 1980 he moved to the California Institute of Technology, and in 1992 he moved yet again to the Swiss Federal Institute of Technology (ETH, Zurich) where he was Professor of Biotechnology until his untimely death. During his 30 years in academia he made seminal contributions to many fields of chemical engineering science, including catalysis and reaction engineering, bioprocess engineering, mathematical modeling of cellular processes, recombinant DNA technology, enzyme engineering, and metabolic engineering. Here we summarize some of his contributions to the engineering of molecular and cellular biocatalysts.

Bailey's growing interest in biological systems in the 1970's and 1980's was no doubt due to his interest in chemical catalysis. So it was no surprise that he moved into the fields of enzyme technology and fermentation/cell cul-

ture. In essence, these areas are catalytic; whether it is in the use of an enzyme to catalyze biotransformations or a cell to produce specific compounds. Although he did not use enzymes synthetically, he did set the stage for research that would lead to several generations of enzymologists, chemists, and biochemical engineers who use enzymes in organic synthesis. By laying the quantitative groundwork for immobilized enzyme systems, and studying their structure and function in synthetically relevant forms, Bailey elevated biocatalysis from a synthetic oddity to a promising technology.

In the mid-1980's, shortly after Bailey started to acquaint himself with the principles and practice of emerging recombinant DNA technologies, he came to realize that the science of metabolism presented a particularly fertile ground for a chemical engineer to apply this new technology. To emphasize the marriage between metabolic biochemistry and chemical engineering science, he coined the term "metabolic engineering." By the mid-1980's he had set his sights squarely on the twin goals of developing the fundamentals of metabolic engineering and identifying interesting applications within this new area of applied science that would highlight its longterm potential to the broader scientific community. This was to remain a dominant theme in his research program for the remainder of his career. His vision also inspired an entire generation of students from his own and other laboratories, many of who continue to evolve the frontiers of metabolic engineering to this day.

## IMMOBILIZED ENZYME TECHNOLOGY

When compared to their chemical counterparts, biocatalysts are exquisitely selective and highly reactive over a broad range of operating conditions. Moreover, whole microbial cells (primarily bacterial and fungal) and their catalytic machinery (e.g., enzymes and metabolic pathways) can accept a wide array of complex molecules as substrates, yielding products with unparalleled chiral (enantio-), positional (regio-), and chemoselectivities. Such high selectivity affords efficient reactions with few byproducts, and ensuring that

Correspondence to: Jonathan S. Dordick and Chaitan Khosla

enzymes can be used in both simple and complex transformations without the need for tedious blocking and deblocking steps that are commonplace in enantio- and regioselective organic synthesis.

### Structure and Function of Enzymes on Solid Supports

Prior to Bailey's arrival in the enzyme technology arena, the vast majority of biocatalysts, even for commercial applications, involved soluble enzymes. Nonetheless, heterogeneous preparations offer a variety of advantages over homogeneous ones for scaled-up biocatalytic operation, including retention of the enzyme in the bioreactor, use of packed-bed operation, and high catalyst loading. Immobilized enzymes, however, also display different structural and functional properties, as well as being subject to substrate diffusional limitations that impact observed enzyme activity and stability.

As early as 1977, Bailey recognized that immobilization could result in operational advantages for biocatalytic systems. For example, glucose oxidase, an enzyme that is rapidly inactivated by  $H_2O_2$  could be stabilized dramatically upon immobilization, even using simple covalent attachment techniques. Various supports were used, including activated carbon, glass, metal oxides, and polymer resins (Bailey and Cho, 1983; Cho and Bailey, 1978, 1979). These early studies were augmented by quantitative evaluation of enzyme systems on solid supports, with direct comparison to soluble systems. Bailey and others found that the observed reaction kinetics could be affected by the nature of the immobilization support. In one of the first applications of electron paramagnetic resonance (EPR) spectroscopy applied to immobilized enzymes, Bailey, along with then graduate student Douglas Clark, demonstrated that immobilization affects the structure and dynamics of the mammalian protease,  $\alpha$ -chymotrypsin (Clark and Bailey, 1984). Structure-function studies were performed, which resulted in a firm understanding of the improvement in enzyme activity upon the use of linkers to attach enzymes to the support rather than direct enzyme attachment. Simultaneous to this study was the ability to quantify the influence of substrate diffusional limitations on heterogeneous enzymatic catalysis (Clark et al., 1985; Dennis et al., 1984). This study, expanded many fold by others on a wide range of commercially relevant enzymes, provided a quantitative foundation for immobilized enzyme technology that is critical today in the food, pharmaceutical, and chemical industries, and in both aqueous and nonaqueous media. Above all else, these studies led to the realization that enzymes [and whole cells (Doran and Bailey, 1986)] were capable of being manipulated simply by the environment that they were used in, and therefore, the enzyme technologist was not limited to the native properties of enzymes for eventual operation.

### Nonaqueous Enzymology—Quantitative and Mechanistic Principles

Enzymatic catalysis in organic solvents has dramatically shaped the emerging use of enzymes in organic synthesis (Dordick, 1992; Klivanov, 1990). Nonaqueous media results in higher substrate solubility, reversal of hydrolytic reactions, modified enzyme specificity, and improved thermostability. Nonaqueous conditions enable one to tap into novel enzyme activities and selectivities heretofore only possible using genetic modifications or complex multistep pathways. Because such an approach is so attractive, enzymatic catalysis in nonaqueous media has undergone rapid expansion, particularly over the past decade. While there is little question that enzymes can function in nonaqueous media, reaction rates are typically quite low and enzymes often have limited stability in such environments. In nearly all cases, the catalytic activity displayed by enzymes in nearly anhydrous (or neat) organic solvents is far lower than in water, as much as five orders of magnitude lower! There is nothing unique about subtilisin, and similarly diminished activity of many enzymes (hydrolases and oxidases) is observed in organic solvents. Nevertheless, there may be nothing inevitable about this decline, and both its underlying causes and effective remedies are emerging.

While Jay Bailey did not study biocatalysis in organic media, nonetheless his prior research yielded the tools necessary to make critical contributions in this area, particularly for studying heterogeneous enzyme preparations quantitatively. For example, in the early 1990's, Affleck et al. (1992) and Xu et al. (1994), using enzyme kinetic, solution thermodynamic, and EPR spectroscopic techniques, uncovered an intriguing correlation between increased activity of enzymes in slightly hydrated organic solvents relative to dry solvents and the polarity of the enzyme's active site. Further quantitative studies involving pressure-based electrostriction resulted in one of the first transition-state mechanistic models of observed enzyme activity in nearly anhydrous environments. Specifically, Michels et al. (1998) used Kirkwood electrostriction analysis to show that the dipole moment of subtilisin's transition state was the same in hexane as in water. Hence, it is not surprising that the enzyme is much less active in the former than the latter, as water can stabilize charge separation in the transition state that gives rise to such a dipole moment. Prior to this result, Khmel'nitsky et al. (1994) discovered that enzymes could be dramatically activated (nearly 4,000-fold) by lyophilizing them in the presence of a nonbuffer salt such as KCl. Based on the work by Michels et al. (1998), it was reasoned that charge stabilization could be achieved by nonbuffer salts, which provides a locally high polarity near the transition state of the enzyme in nonpolar organic media, even in the absence of added water. Indeed, the 1990's provided a wealth of techniques to activate enzymes for use in organic media. Some of the prominent ones include the addition of polyols (Adlercreutz, 1993), crown ethers (Engbersen et al., 1996), transition-state analogs (Slade and Vulfson, 1998), and substrates and substrate mimics (Bracco et al., 1990; Rich and



Dordick, 1997). Beginning with a fundamental picture of enzymes in the nonaqueous milieu, the field of nonaqueous enzymology has recently seen cases where enzymes function in organic media as well as they do in water; again a striking realization that biocatalysis can be tailored by manipulating an enzyme's heterogeneous environment. Such a result can be traced back to Jay Bailey and his initial work with glucose oxidase and  $\alpha$ -chymotrypsin on solid supports.

Activated biocatalyst formulations have begun to impact society in terms of human health (better drugs), the environment (more efficient and selective synthetic conditions), and industry (new routes to existing and novel chemicals and materials). For example, salt-activated thermolysin catalyzes the regioselective acylation of paclitaxel (taxol) in a synthetic scheme to produce a water-soluble prodrug (paclitaxel 2'-adipic acid), which may result in an easier mode of delivery of this anticancer compound (Khmelnitsky et al., 1997). The adipic acid derivative of taxol is  $\approx 1700$ -fold more soluble in water than the native drug. In addition to the direct impact on taxol prodrug synthesis, activated biocatalysts are beginning to find applications in the synthesis of chiral molecules and in the resolution of optical isomers, both critical in the preparation of new pharmaceuticals.

Nonaqueous enzymology is quickly maturing, and combined with directed evolution (Affholter and Arnold, 1999) and gene-shuffling technologies (Powell et al., 2001), biocatalysts with tailored and controllable activities, stabilities, and selectivities will yield still better commercial catalysts. The ability to use enzymes in organic media has also led to the bridging of the materials and biological worlds, wherein enzymes and other proteins can be incorporated into polymers (organic and inorganic) to provide structure and function to the material (Wang et al., 1997). Opportunities in drug discovery, through protein chip technologies (Kim et al., 2001), and nanotechnology (Graff et al., 2001), through smart materials, will have a dominating impact in society for years to come. All of these advances can be linked to the fundamental immobilized enzyme studies initiated by Jay.

## MANIPULATION OF CELLULAR FUNCTION

Three main reasons make metabolism a particularly interesting target for manipulation by a chemical engineer. First, in its most basic form, metabolism represents a large set (network) of chemical transformations operating in series/parallel in one pot. This has a direct analogy with, say, the chemistry in a naphtha cracker, the analysis of which is widely regarded as the fountainhead of modern chemical reaction engineering. Jay Bailey was only too aware of this, since his graduate training and early independent research focused on the theoretical and experimental analysis of complex reaction networks. Second, most practical limitations of cells as biocatalysts are due to metabolic constraints or defects. It therefore follows that the enhancement of existing biotechnological processes as well as the development of new ones rests heavily on a solid appreciation of the nuances of the underlying metabolism. Finally, the synthetic

capabilities of metabolism (especially anabolic metabolism) are unparalleled. Thus, the ability to manipulate such powerful chemistry should be able to afford a virtually unlimited spectrum of new products. Since Bailey was especially interested in the practical applications of academic science, the latter two arguments also appealed to him.

What does a chemical engineer need to learn to tackle metabolic problems? Bailey thought long and hard about this right from the beginning. He was particularly aware that metabolic engineering would be unable to sustain itself as a burgeoning field if it simply represented a collection of interesting anecdotal examples involving the manipulation of cellular function. Indeed, the title of his landmark article in *Science* (1991) is testament to his constant desire for a strong foundation based on rigorously tested principles and cutting-edge tools. Examples of such principles and tools include:

1. *The chemical logic of metabolism.* While not apparent at first glance, the chemical logic of metabolism provides an excellent basis for organizing an overwhelmingly large body of information into a "sensible" network of chemical transformations. For example, an understanding of how and why energy is stored in the form of phosphoanhydride linkages in a cell allows one to appreciate the logic of many phosphorylation-dephosphorylation reactions that occur during central carbon metabolism (glycolysis and TCA cycle). Likewise, by understanding the power of the aldol reaction, one can readily recognize how the six carbon atoms of glucose might be scrambled to yield the carbon backbone of each of the 20 amino acids in a cell. When viewed through the eyes of chemical structure and reactivity, metabolic stoichiometry (and to some extent, even metabolic kinetics) are no longer "dry facts," and a rational approach can be taken toward strategic decisions regarding which reactions to manipulate toward a metabolic engineering objective.
2. *The biological logic of metabolism.* While many aspects of metabolic reaction networks are conceptually related to networks associated with other problems in chemical reaction engineering, it is the biological control of metabolism that sets this problem area apart. Concepts such as positive and negative regulation of enzyme activity at a transcriptional and/or posttranscriptional level must be understood well by chemical engineers before they can conduct a meaningful analysis of a metabolic problem. The fact that a living cell has devised a dazzling range of circuits for metabolic control, most of which have no parallel in nonliving systems, does not make the challenge any easier for the chemical engineer, but a background in molecular biology can be invaluable in this regard. The development of effective ways to translate the molecular biologist's cartoon representations of metabolic control into quantitative models was a recurrent theme in Jay Bailey's research (see, for example,

Lee and Bailey, 1984), and remains a major challenge in metabolic engineering to this day.

3. **Chemical tools.** Tools of modern analytical chemistry are especially valuable to a metabolic engineer in his analysis of perturbation-response experiments. Often, due to the highly coupled and nonlinear connections that exist within a metabolic network, it is not possible to intuitively predict the outcome of a metabolic perturbation achieved via genetic or environmental manipulation. In conjunction with steady-state or pulse-chase experiments, analytical tools such as GC/LC-MS, UV spectroscopy, fluorescence spectroscopy, and NMR spectroscopy allow one to monitor changes in concentrations and fluxes of metabolic intermediates (Bailey et al., 1987). Bailey saw the power of each one of these tools, as well as many others, and learned to not only apply them effectively to metabolic problems but also to innovate with them.
4. **Biological tools.** Above all else, the toolbox of modern molecular biology has been the driving force for the emergence of metabolic engineering science and applications. Molecular biology has provided an unparalleled ability to manipulate genes individually or combinatorially, and also to monitor structural and dynamic changes in cellular macromolecules such as DNA, RNA, and proteins. At a time when many chemical engineers were wary of the impact of molecular biology tools on their field, Bailey was one of the first chemical engineers to wholeheartedly embrace this toolbox and to put virtually every tool to use in his attack on metabolic problems. More recently, he recognized the value of genomic and proteomic tools in extending the reach of molecular biology to multi-variable analysis of metabolism.
5. **Mathematical tools.** Although metabolism is a descriptive science, metabolic engineering is most effective when undertaken on the basis of a quantitative framework. Herein lies a particularly vexing challenge—although the formalism of metabolic control theory is adequately robust to describe most problems in metabolic engineering, the complexity of most practical applications in metabolic engineering preclude rigorous application of this theory. Major (often drastic) approximations must be made that limit the generality of mathematical models. Jay Bailey experimented with the entire gamut of mathematical tools for quantitatively describing and analyzing metabolic processes, ranging from rigorous coupled nonlinear differential equations, to cybernetic models based on “soft” axioms. This was a particularly strong passion of his during his last years. His final essay titled “Complex biology with no parameters” provided advice as well as exhortation to present-day metabolic engineers (Bailey, 2001).

In addition to developing and refining metabolic engineering principles and tools, Jay Bailey also made contributions to many important problems in applied metabolic engineering. Examples of problems he studied included pH

homeostasis in microorganisms, central carbon metabolism in *E. coli* and yeast, oxidative (especially microaerobic) metabolism, posttranslational modifications of proteins in mammalian cell culture, and product-oriented niche metabolism such as solventogenic fermentation processes. Within his exceptionally broad research program, each of these problems was marked by a desire to understand, control, and manipulate the metabolic property of interest. Bailey’s choice of problems and his approach to a solution was characteristic, and had an enormous influence on the field. His work has been widely cited and will almost certainly be studied carefully by future generations of metabolic engineers.

## DRUG DISCOVERY

Because of the complexity of many proven biologically active compounds, traditional drug discovery methods have begun to be supplemented by nontraditional methodologies that focus on either rational or combinatorial techniques for drug discovery and development. The former requires extensive knowledge about a biological target (e.g., a receptor’s binding site or an enzyme’s active site), and utilizes computational chemistry and molecular modeling to design chemical structures that may illicit a biological response. Often, such molecular targets are either not available, or their structures are not available in sufficient resolution to provide suitable targets for molecular modeling. In these cases, drug discovery has recently turned to combinatorial techniques for new lead discovery and development. Within the past decade there has been a major growth in the application of engineered enzymes and whole cells to problems in drug discovery. Two such examples are described below.

## Combinatorial Biology

Secondary metabolism refers to metabolic processes in a cell that occur during post-exponential (often stationary) phase of growth, and are therefore unnecessary for the survival or reproductive capacity of a growing cell. Many microorganisms and plants produce structurally complex natural products as secondary metabolites; some of these have exquisite biological activities that have led to their exploitation as antibiotics, anticancer agents, or other pharmacologically useful agents. Examples of well known natural products synthesized as secondary metabolites include the antibiotics penicillin, streptomycin, erythromycin, tetracycline, vancomycin, the anticancer agents adriamycin, taxol and vinblastine, the cholesterol lowering agents compactin and lovastatin, and the immunosuppressants cyclosporin, FK506, and rapamycin.

Although it has been recognized for many decades now that, notwithstanding structural distinctions, the biosynthetic pathways of many of these natural products are related, an understanding of the associated catalytic mechanisms is only just beginning to emerge. Hand in hand with these fundamental insights, the application of protein engi-

neering principles to natural product biosynthesis has resulted in the emergence of a new field, often referred to as combinatorial biosynthesis, where the structure of a natural product is systematically manipulated by genetic manipulation of the biosynthetic enzymes. Combinatorial biosynthesis has yielded numerous new "unnatural" natural products over the past 10 years, and can be used to optimize the properties of existing and emerging bioactive natural products (Rodríguez and McDaniel, 2001). In addition, it could also be used to construct new natural product libraries for drug discovery. In particular, the enzymes responsible for biosynthesis of four major classes of natural products—polyketides, nonribosomal peptides, isoprenoids, and deoxy-sugars and related aminocyclitols—are emerging as especially fertile targets for genetic and chemo-biosynthetic manipulation (Cane et al., 1998). The first products from such biosynthetic engineering efforts are already entering clinical trials. As more such engineered metabolites emerge from discovery into development, the molecular vision that Jay Bailey gave to the field of biochemical engineering will undoubtedly be realized.

### Combinatorial Biocatalysis for Lead Compound Optimization

Rapid developments in genomics, proteomics, and combinatorial chemistry have reshaped the field of drug discovery, providing new drug targets for selective screens and new compounds to be tested in those screens. While combinatorial methods have given rise to large libraries of compounds, typically these compounds result in improved lead candidates that must undergo further transformations by conventional medicinal chemistry to yield new drug candidates. High-throughput combinatorial methodologies have not impacted lead optimization nearly as much as they have lead discovery, mainly because of the highly selective, intricate chemistries often required to optimize lead compounds. This is particularly challenging for optimization of natural products or complex synthetic leads, the latter often coming from initial combinatorial synthesis and high-throughput screening.

Nature's most potent molecules are produced by enzymecatalyzed reactions coupled with natural selection of those products with optimal biological activity. Combinatorial biocatalysis harnesses the natural diversity of enzymatic reactions for the synthesis of organic compound libraries to generate biologically active compounds, which encompass a wide array of chemistries and structures (Michels et al., 1998). Combinatorial biocatalysis is focused on both biocatalytic transformation and iterative synthesis. The more complex the lead compound, the more iterations that are possible and the larger the library of derivatives. Hence, thousands of derivatives of the original lead compound can be produced using combinatorial biocatalysis. Initial development of this technology focused on the generation of solution-phase combinatorial libraries, including those from synthetic precursors (e.g., dibenzyl 1,2-phenylenedioxydi-

acetate as bis-amide derivatives) (Adamczyk et al., 1997, as well as natural products such as flavonoids (e.g., bergienin), polyketides (e.g., doxorubicin and erythromycin) (Altreuter et al., 2002), nucleosides (e.g., adenosine), and diterpenoids (e.g., paclitaxel) (Khmel'nitsky et al., 1997) have been generated in solution using enzymes or whole cells and their extracts. Together with combinatorial biology, the ability to tap into even a small part of nature's vast repertoire of biocatalytic machinery is now possible.

### FINAL THOUGHTS

Jay Bailey trained as a chemical engineer during the heyday of catalytic reaction engineering as practiced in the petrochemical industry. He foresaw the impact that enzymes and cells were likely to have on the development of new products and processes, and dedicated himself to the visionary goal of educating new students of chemical engineering about the power of modern biology in this endeavor. Just a few short decades ago, not only were chemical engineers ignorant of biology, but it was difficult to make a compelling case for the relevance of biology to our discipline. The fact that catalysis and biocatalysis have merged so intimately in such a short timespan is a wonderful testament to Jay Bailey's impact on chemical engineering. Now that biology is becoming a quantitative discipline, the impact of Jay Bailey extends to the biomolecular sciences, including the interface areas of genomics and proteomics, nanobiotechnology, and high-throughput drug discovery. Thus, Bailey will impact biology in years to come as much as he impacted biochemical engineering in the past three decades. In 2050, as chemical engineers prepare to inaugurate the first zero-emission manufacturing facility that converts atmospheric carbon dioxide into automotive fuel using an engineered multienzyme system, microorganism, or plant as a catalyst, they will trace the roots of their accomplishments back to Jay Bailey's vision for molecular bioengineering in very much the same way as modern aircraft designers recognize the contributions of von Karman in the area of jet propulsion.

### References

- Adamczyk M, Gebler JC, Grote J. 1997. The utility of enzymes in generating molecular diversity. Lipase mediated amidation of polybenzyl esters. *Bioorg Med Chem Lett* 8:1027-1030.
- Adlercreutz P. 1993. Activation of enzymes in organic media at low water activity by polyols and saccharides. *Biochim Biophys Acta* 1163: 144-148.
- Affholter J, Arnold FH. 1999. Engineering a revolution. *Chemtech* 9: 34-39.
- Affleck R, Xu Z-F, Suzawa V, Foelt K, Clark DS, Dordick JS. 1992. Enzymatic catalysis and dynamics in low-water environments. *Proc Natl Acad Sci USA* 89:1100-1104.
- Altreuter DA, Dordick JS, Clark DS. 2002. Nonaqueous biocatalytic synthesis of new cytotoxic doxorubicin derivatives: Exploiting unexpected differences in the regioselectivity of salt-activated and solubilized subtilisin. *J Am Chem Soc* 124:1871-1876.

- Bailey JE. 1991. Toward a science of metabolic engineering. *Science* 251: 1668-1675.
- Bailey JE. 2001. Complex biology with no parameters. *Nat Biotechnol* 19:503-504.
- Bailey JE, Axe DD, Doran PM, Galazzo JL, Reardon KF, Seressiotis A, Shanks JV. 1987. Redirection of cellular metabolism. Analysis and synthesis. *Ann NY Acad Sci* 506:1-23.
- Bailey JE, Cho YK. 1983. Immobilization of glucoamylase and glucose oxidase in activated carbon: Effects of particle size and immobilization conditions on enzyme activity and effectiveness. *Biotechnol Bioeng* 25:1923-1935.
- Braco L, Dabulis K, Klibanov AM. 1990. Production of abiotic receptors by molecular imprinting of proteins. *Proc Natl Acad Sci USA* 87: 274-277.
- Cane DE, Walsh CT, Khosla C. 1998. Harnessing the biosynthetic code: Combinations, permutations, and mutations. *Science* 282:63-68.
- Cho YK, Bailey JE. 1978. Immobilization of enzymes on activated carbon: Properties of immobilized glucoamylase, glucose oxidase, and gluconolactonase. *Biotechnol Bioeng* 20:1651-1665.
- Cho YK, Bailey JE. 1979. Immobilization of enzymes on activated carbon: Selection and preparation of the carbon support. *Biotechnol Bioeng* 21:461-476.
- Clark DS, Bailey JE. 1984. Characterization of heterogeneous immobilized enzyme subpopulations using EPR spectroscopy. *Biotechnol Bioeng* 26:231-238.
- Clark DS, Do DD, Bailey JE. 1985. A mathematical model for restricted diffusion effects on macromolecule impregnation in porous supports. *Biotechnol Bioeng* 27:208-213.
- Dennis KE, Cho YK, Clark DS, Park YH, Bailey JE. 1984. Immobilization of enzymes in porous supports: Effects of support-enzyme solution contacting. *Biotechnol Bioeng* 26:892-900.
- Doran PM, Bailey JE. 1986. Effects of immobilization on growth, fermentation properties, and macromolecular composition of *Saccharomyces cerevisiae* attached to gelatin. *Biotechnol Bioeng* 28:73-87.
- Dordick JS. 1992. Designing enzymes for use in organic solvents. *Biotechnol Prog* 8:259-267.
- Engbersen JFJ, Broos J, Verboom W, Reinhoudt DN. 1996. Effects of crown ethers and small amounts of cosolvent on the activity and enantioselectivity of  $\alpha$ -chymotrypsin in organic solvents. *Pure & Appl Chem* 68:2171-2178.
- Graff A, Winterhalter M, Meier W. 2001. Nanoreactors from polymer-stabilized liposomes. *Langmuir* 17:919-923.
- Khmelnitsky YL, Welch SH, Clark DS, Dordick JS. 1994. Salts dramatically enhance activity of enzymes suspended in organic solvents. *J Am Chem Soc* 116:2647.
- Khmelnitsky YL, Budde C, Arnold JM, Usyatinsky A, Clark DS, Dordick JS. 1997. Synthesis of water-soluble paclitaxel derivatives by enzymatic acylation. *J Am Chem Soc* 119:11554-11555.
- Kim Y-D, Park C-B, Clark DS. 2001. Stable sol gel microstructured and microfluidic networks for protein patterning. *Biotechnol Bioeng* 73: 331-337.
- Klibanov AM. 1990. Asymmetric transformations catalyzed by enzymes in organic solvents. *Acc Chem Res* 23:114.
- Lee SB, Bailey JE. 1984. Analysis of growth rate effects on productivity of recombinant *Escherichia coli* populations using molecular mechanism models. *Biotechnol Bioeng* 26:66-73.
- Michels PC, Dordick JS, Clark DS. 1997. Dipole formation and solvent electrostriction in subtilisin catalysis. *J Am Chem Soc* 119:9331.
- Michels PC, Khmelnitsky YL, Dordick JS, Clark DS. 1998. Combinatorial biocatalysis: A natural approach to drug discovery. *Trends Biotechnol* 16:210-215.
- Powell KA, Ramer SW, Del Cardayre SB, Stemmer WPC, Tobin MB, Longchamp PF, Huisman GW. 2001. Directed evolution and biocatalysis. *Angew Chem Int Ed* 40:3948-3959.
- Rich JO, Dordick JS. 1997. Controlling subtilisin activity and selectivity in organic media by imprinting with nucleophilic substrates. *J Am Chem Soc* 119:3245.
- Rodriguez E, McDaniel R. 2001. Combinatorial biosynthesis of antimicrobials and other natural products. *Curr Opin Microbiol* 4:526-534.
- Slade CJ, Vulfson EN. 1998. Induction of catalytic activity in proteins by lyophilization in the presence of a transition state analogue. *Biotechnol Bioeng* 57, 211-215.
- Wang P, Sergeeva M, Lim L, Dordick JS. 1997. Biocatalytic plastics as active and stable materials for biotransformations. *Nature Biotechnology* 15:789.
- Xu Z-F, Affleck R, Wangkar P, Suzawa V, Dordick JS, Clark DS. 1994. Transition state stabilization of subtilisins in organic media. *Biotechnol Bioeng* 43:515.



Ecole polytechnique fédérale de Zurich  
Politecnico federale di Zurigo  
Swiss Federal Institute of Technology Zurich

Professor Dr. James E. Bailey  
Institute of Biotechnology

ETH Zürich  
CH 8093 Zürich,  
Switzerland

Tel.: 41-1-633 31 70  
41-1-633 34 03  
FAX: 41-1-633 10 51

Professor Bernhard O. Palsson  
Department of Bioengineering  
University of California,  
San Diego  
La Jolla, CA 92093  
USA

FAX: 001 619 822 0240

Zurich, 8 July 1999

Dear Bernhard,

I look forward to seeing you again at the Biochemical Engineering Conference in Salt Lake City. In anticipation of that, I am writing with some questions about your genomics-flux balance approach.

Your recent paper in *Biotechnology Progress* is a very effective synopsis of general aspects of genomics and future challenges which I think is very useful for all bioengineers to see. The rest of the paper, presenting a number of case studies and results which are based on flux analysis, is harder for me to understand.

As I can infer from the paper, you are somehow calculating how flux distributions and specific growth rates change when certain genes are not expressed. This calculation seems to rest on an optimization calculation, which I guess must be a linear programming problem, in which the fluxes are the decision variables and making cell mass is the objective function. How does this work exactly? What is the optimization problem formulation? What are the constraints besides the metabolite mass balances? What metabolites are assumed to enter and exit the cell, and what determines those rates? And, finally, a major question, is how do you arrive at conclusions about, say, changes in specific growth rates, given a stoichiometric network? How do you get from stoichiometry to rates?

The matter of enzymes added to those implied by the genome sequence (47 of 587 according to your BP paper) is also a question. Could you please send me a list of these added steps?

I hope that you might take some time in your presentation in San Diego to explain some of these points in some detail. This method, if it is correct and if it works, is a very major advance, because it gives a formalism for determining growth rates and pathway rates without any knowledge of any kinetics. I must say it is little hard to believe that this can be done, but maybe you have made a breakthrough.

Any feedback you could give me before the meeting would be much appreciated.  
Thanks very much. See you soon.

Best wishes,

A handwritten signature in black ink, appearing to be 'Jay' or similar, written in a cursive style.

# Robustness Analysis of the *Escherichia coli* Metabolic Network

Jeremy S. Edwards<sup>1</sup> and Bernhard O. Palsson\*

Department of Bioengineering, University of California-San Diego, La Jolla, California 92093-0412

Genomic, biochemical, and strain-specific data can be assembled to define an *in silico* representation of the metabolic network for a select group of single cellular organisms. Flux-balance analysis and phenotypic phase planes derived therefrom have been developed and applied to analyze the metabolic capabilities and characteristics of *Escherichia coli* K-12. These analyses have shown the existence of seven essential reactions in the central metabolic pathways (glycolysis, pentose phosphate pathway, tricarboxylic acid cycle) for the growth in glucose minimal media. The corresponding seven gene products can be grouped into three categories: (1) pentose phosphate pathway genes, (2) three-carbon glycolytic genes, and (3) tricarboxylic acid cycle genes. Here we develop a procedure that calculates the sensitivity of optimal cellular growth to altered flux levels of these essential gene products. The results indicate that the *E. coli* metabolic network is robust with respect to the flux levels of these enzymes. The metabolic flux in the transketolase and the tricarboxylic acid cycle reactions can be reduced to 15% and 19%, respectively, of the optimal value without significantly influencing the optimal growth flux. The metabolic network also exhibited robustness with respect to the ribose-5-phosphate isomerase, and the ribose-5-phosphate isomerase flux was reduced to 28% of the optimal value without significantly affecting the optimal growth flux. The metabolic network exhibited limited robustness to the three-carbon glycolytic fluxes both increased and decreased. The development presented another dimension to the use of FBA to study the capabilities of metabolic networks.

## Introduction

Genome sequencing and bioinformatics are beginning to reveal the complete set of molecular components involved in cellular activities. Furthermore, it is also clear that the integrated function of biological systems involves complex interactions among the components that have been identified through bioinformatics and genomics. Importantly, the properties of complex systems cannot be predicted simply on the basis of the complete description of their components, and the emergent properties of biological systems need to be studied (1, 2). To understand the complexity inherent in cellular networks, approaches that focus on the systemic properties of the network are required. The focus of such research represents a departure from the classical reductionist approach to the integrated approach (3) to understanding the interrelatedness of gene function and the role of each gene in the context of multigenetic cellular functions or genetic circuits (4, 5).

The engineering approach to analysis and design is to have a mathematical or computer model, e.g., a dynamic simulator, of a cellular process that is based on fundamental physicochemical laws and principles. There has been a long history of mathematical modeling of metabolic systems, which dates back to the mid 1960s. With the availability of analogue computers and the knowledge of metabolic regulation, dynamic simulations

of simple metabolic and genetic control loops appeared (6). The dynamic stability of such control loops became a focus of attention (7, 8), given the experimental observations of oscillatory dynamics in yeast glycolysis (9).

The systemic nature of metabolic function was apparent, and so was its complexity. However, the availability of enzyme kinetic information was fragmented, and attention turned to developing methods that could shed light on the relative importance of various metabolic events. Methods for sensitivity analysis of metabolic regulation began in the 1960s (10) and continued into the 1970s (11, 12). The results of these undertakings were biochemical systems theory (BST) and metabolic control analysis (MCA), and some useful results have been obtained using these approaches (13).

Establishing complete kinetic models of cellular metabolism became a scientific goal, whose intended use was to elucidate the systemic behavior of metabolic networks. Because of its simplicity, the human red blood cell represented the best opportunity to achieve this goal. Early metabolic models of human red blood cell metabolism appeared in the 1970s (14) and continued throughout the 1980s and 1990s (14-16). Insights into the functioning of this cell have resulted from these analyses (11, 17, 18). Although interesting in their own right, studies of red cell metabolism are not directly useful for organisms of industrial importance.

While the ultimate goal is the development of dynamic models for the complete simulation of metabolic systems, the success of such approaches has been severely hampered by the current lack of kinetic information on the dynamics and regulation of metabolic reactions. However,

\* Ph: (619) 534-5668. Fax: (619) 822-0240. Email: palsson@ucsd.edu.

<sup>1</sup> Current address: Department of Genetics, Harvard Medical School, Boston, MA 02115.

in the absence of kinetic information it is still possible to accurately assess the theoretical capabilities and operative modes of metabolic systems using metabolic flux balance analysis (FBA) (5, 19–23). FBA is based on the fundamental physicochemical constraints on metabolic networks. FBA only requires information regarding the stoichiometry of metabolic pathways and the metabolic demands; furthermore, FBA can incorporate additional information when it is available. FBA is particularly applicable for post-genomic analysis, because the stoichiometric parameters can be defined from the annotated genome sequence (21).

In a previous article, we have examined the capability of *in silico* mutant *E. coli* metabolic networks to support growth and compared the results to the wildtype. By using computer simulations, it was determined that seven metabolic reactions were essential for the aerobic growth of *E. coli* in glucose minimal media (24). The remaining reactions were determined to be nonessential, since the metabolic network maintained the capability to bypass simulated metabolic defects, often with little or no effect on the *in silico* maximal biomass yield. In this article we will further examine the essential metabolic reactions by examining the metabolic consequences of reduced metabolic flux carrying capacity in the essential reactions. The results indicate the redundancy and robustness in the function of the respective metabolic reactions in the metabolic network by examining the sensitivity of the objective function to the quantitative flux levels. The sensitivity analysis can provide information regarding the experimental measurements that are likely to provide the most information toward quantitatively describing the metabolic network and can be used for *in silico* experimental design and assessing the value of the *in silico* predictions.

### Describing Metabolic Systems

A metabolic network is a collection of enzymatic reactions that serve to biochemically process metabolites within the cell and transport processes that convert extracellular metabolites to intracellular metabolites and vice versa. To quantitatively describe metabolic networks, dynamic mass balances are written for each metabolite in the network, generating a system of ordinary differential equations that describe the transient behavior of metabolite concentrations:

$$\frac{dX_i}{dt} = \sum_j S_{ij} v_j \quad (1)$$

where  $v_j$  corresponds to the  $j$ th metabolic flux,  $X_i$  represents the  $i$ th metabolite, and the stoichiometric coefficient  $S_{ij}$  stands for the number of moles of metabolite  $i$  formed (or consumed) in reaction  $j$ . Equation 1 is particularly difficult to solve since the metabolic fluxes are often nonlinear functions of the metabolite concentrations, as well as a set of kinetic parameters that are difficult to measure or estimate. The complexity associated with estimating the functional relation between the metabolic fluxes and the metabolite concentrations and the associated kinetic parameters has hampered the quantitative analysis of metabolic networks.

### Constraining Metabolic Functions

Given the complexities associated with quantitative analysis of metabolic systems based on kinetic characterization of the components, we have utilized a conceptually different approach to the analysis of metabolic networks. First, we defined fundamental physicochemical

constraints to which the metabolic network is constrained. Then, the metabolic capabilities were assessed subject to the imposed constraints. The capabilities are analyzed under the steady state assumption. It should be noted that steady state analysis is applicable to some aspects of metabolism; however, the approach will not be appropriate for studying all cellular processes, such as the cell cycle or signal transduction. Herein, we are interested in metabolic processes and their relation to cellular growth; thus the characteristic time of the processes is about an hour. Metabolic transients within the cell typically occur with time constants on the order of seconds to minutes (25); thus under our "window of observation" the metabolic network is essentially in a steady state and the steady-state analysis will be appropriate. The steady-state mass, energy, and redox balance constraints are imposed by simplifying eq 1:

$$S \cdot v = 0 \quad (2)$$

where  $S$  is the stoichiometric matrix and  $v$  is the flux vector. While the system is closed to the passage of certain metabolites, others are allowed to enter or exit the system via exchange fluxes (or pseudoreactions (26)). These fluxes do not represent biochemical conversions or transport processes such as those of internal fluxes but can be thought of as representing the inputs and outputs to the system. For example, the demand on a metabolite for further processing or incorporation into cellular biomass creates an exchange flux on the internal cellular metabolite. Thus, a distinction is made between internal and external metabolites in the system, therefore closing the material balance to all metabolites as indicated by eq 2.

To complete the *in silico* representation of the metabolic network we included the constraints on the individual metabolic reaction fluxes due to reaction thermodynamics and the input/output characteristics of the network. All reversible metabolic reactions were assumed to have the capability to carry any metabolic flux (i.e.,  $-\infty \leq v_i \leq \infty$ ; where  $v_i$  is the flux in reversible reactions), whereas irreversible metabolic reactions fluxes were restricted to be positive (i.e.,  $0 \leq v_i < \infty$ ; where  $v_i$  is the flux in irreversible reactions). Although constraints on the internal fluxes were defined as infinite, the magnitude of each flux in the optimal solution was examined and compared to measured fluxes (27, 28). The reversibility of each reaction in the metabolic network was determined case by case on the basis of the literature and compared to the EcoCyc database (29). The metabolic enzymes identified in the complete *E. coli* K12 genome sequence and the online databases (29–31) were used to reconstruct the metabolic network (see supplementary information at [http://gcrd.ucsd.edu/supplementary\\_data/BP2000/main.htm](http://gcrd.ucsd.edu/supplementary_data/BP2000/main.htm)). It should be noted that there are instances where the same enzyme can catalyze multiple reactions (e.g., different substrates or cofactors), and this situation was considered by including all reactions catalyzed by an enzyme as a separate column in the stoichiometric matrix. The details of this metabolic reconstruction have been described elsewhere (24). Additionally, constraints were placed on the exchange fluxes to indicate the environmental conditions. For example, metabolites not available to the cell are constrained to not enter the cell:  $-\infty \leq b_i \leq 0$ , where  $b_i$  (influx defined as positive) is the exchange flux for a metabolite not available in the simulated environment. It should be noted that all metabolites that have the capability to leave the cell always had unconstrained metabolic fluxes



in the net outward direction, whereas the influx constraints were defined by the simulated environmental conditions. For the analysis herein, the exchange flux for inorganic phosphate, ammonia, carbon dioxide, oxygen, sulfate, potassium, and sodium were unconstrained, whereas the uptake of the carbon source was constrained as specified.

### Demands on the Metabolic Network

Under changing substrate/supply conditions metabolic networks are continuously faced with a balanced set of biosynthetic demands (i.e., production of amino acids, nucleotides, phospholipids, as well as energy and redox potential). Effectively this means that the network must generate a balanced set of metabolites that are used to produce biomass. The biosynthetic demands for growth were determined from the biomass composition of *E. coli* (32, 33), and a metabolic flux, defined as the growth flux ( $V_{\text{growth}}$ ), utilizes the biosynthetic precursors in the appropriate ratios so as to generate biomass:

$$\sum_{i=1}^n d_i \cdot X_i \xrightarrow{V_{\text{growth}}} \text{Biomass}$$

where  $d_i$  (mmol  $\cdot$  g-dry weight (DW) $^{-1}$ ) is the *E. coli* biomass composition of metabolite  $i$ . One gram of biomass is produced per unit flux in the growth flux,  $V_{\text{growth}}$ , and if the fluxes are represented with a basis of 1 g-DW  $\cdot$  h $^{-1}$  (23), the growth flux is equivalent to the growth rate. The biomass composition is not constant but depends on the growth rate and the growth conditions (33). However, we have assumed that the biomass composition is constant since it has been shown that the optimal solution is not sensitive to the biomass composition (34), and this observation is also true for our system.

In addition to the biosynthetic demands on the metabolic network, we have also imposed maintenance requirements on the metabolic system. The maintenance requirements included were for growth-associated and non-growth-associated maintenance. We imposed a growth-associated maintenance of 23 mmol ATP  $\cdot$  g-DW $^{-1}$  and a non-growth-associated maintenance of 5.87 mmol ATP  $\cdot$  g-DW $^{-1}$   $\cdot$  h $^{-1}$  (33).

### Exploring the Metabolic Capabilities

The constraints on the metabolic network define the boundaries within which the metabolic system must operate. The mass, energy, and redox balance constraints are imposed by the linear homogeneous set of equations (eq 2). The nullspace of the stoichiometric matrix,  $S$ , contains all flux vectors that satisfy the mass, energy, and redox balance constraints (30). However, there are additional physicochemical constraints on the metabolic network, such as the thermodynamic constraints and the capacity constraints on the exchange fluxes, which are enforced by linear inequalities. The simultaneous enforcement of all the metabolic constraints defines a region, the *feasible set*, that contains all feasible metabolic flux vectors. The feasible set is not a vector space as is the nullspace, as a result of the linear inequality constraints. Importantly, the feasible set defines the metabolic capabilities of the system. The performance capabilities of any metabolic network reside in the feasible set. In fact, the answer to any question related to the general structure and fitness of the network lies within this region. While the feasible set offers a convenient way of defining metabolic capabilities, the question arises, how do we best explore the specific functions of a metabolic network?

One approach that has been used to explore the relationship between the metabolic genotype and phenotype for a number of organisms is linear optimization (19, 21, 22, 37). Linear optimization was used to determine the optimal flux distributions within a network so as to maximize/minimize a particular objective function. A linear programming problem is defined as follows, where a linear objective function is maximized or minimized subject to a series of linear equality and inequality constraints:

$$\begin{aligned} \text{Maximize/Minimize } Z &= c_j v_j \\ \text{subject to } S_{ij} v_j &= 0, \alpha_j \leq v_j \leq \beta_j \end{aligned} \quad (5)$$

The linear programming formalism is analogous to the system of linear equalities/inequalities that form the constraints on the metabolic network. The objective function,  $Z$ , is defined by assigning the appropriate values to the  $c$  vector; herein, the  $c$  vector was taken as the unit vector in the direction of the growth flux. We used the reduced costs from the linear programming solution to identify alternate optimal solutions. In metabolic engineering applications, the objective function can correspond to a number of diverse objectives, such as maximizing energy or metabolite production (20). However, regardless of the objective function the optimal solution will lie within the feasible set that is defined by the physicochemical constraints placed on the system.

The utilization of linear programming to examine metabolic networks defines the optimal flux vector that maximizes (or minimizes) an objective function and satisfies the entire set of constraints. The utilization of design related objectives (such as maximizing the production of an amino acid) can be used to guide genetic engineering of a strain for metabolite overproduction. Herein, we have employed a physiologically realistic objective, the maximization of the growth flux. We have assumed that the cell has evolved the regulatory mechanisms to operate optimally within the feasible set. The feasible set defines the capabilities of the metabolic network, and all metabolic flux vectors within the feasible set satisfy the imposed physicochemical constraints. Therefore, theoretically all flux vectors within the feasible set can be reached by adjusting the enzyme kinetic parameters and gene regulation. The enzyme kinetics and gene regulation constraints on the metabolic system will be referred to as system specific constraints. We assume that the cell has found the optimal set of system specific constraints through the course of evolution, and we attempt to find the same solution using linear programming. The assumption has been experimentally examined under a limited number of conditions, and under defined conditions with a single carbon source, the experimental data is consistent with the optimal utilization of the metabolic network (27).

### Phenotype Phase Plane Analysis

Flux balance analysis can be used to examine the metabolic network in detail. Optimal solutions to the linear programming problem will then lie on a vertex of the feasible set, which is a polyhedron (38). All the metabolic flux vectors (or metabolic phenotypes) attainable from a defined metabolic genotype are mathematically confined to the feasible set. Linear programming was used to search through the feasible set for a solution that maximizes the growth flux. Experimental data for the growth of *E. coli* under nutritionally rich growth conditions (i.e., cell is not starved for phosphate, nitrogen,

etc.) is consistent with the optimal utilization of the metabolic network (27); thus, defining the growth flux as the objective function produces physiologically meaningful results. However, the optimal flux distribution is only meaningful when interpreted in terms of the specific environmental conditions. Therefore, phenotype phase planes (39) have been developed to define the range of optimal flux vectors and how the optimal flux vector is dependent on the environmental conditions.

The methodology for defining PhPPs has been described (39). We will now briefly describe the construction of PhPPs. Two metabolic fluxes can form two axes on an ( $x, y$ )-plane (these metabolic fluxes were two unit vectors in  $R^n$ ). The optimal metabolic flux distribution is calculated for all points in this plane. In other words, the maximum value of the objective function is found as the position of the hyperplanes that bound the feasible set in the respective directions is moved. It has been determined that there are a finite number of fundamentally different optimal metabolic flux distributions (or basis solutions in linear programming terminology) present in such a plane. The demarcations on the phase plane were defined by a shadow price (LP dual variable) analysis (40). This procedure leads to the definition of distinct regions, or "phases", in the plane, in which the optimal use of the metabolic network is fundamentally different, corresponding to different optimal phenotypes.

### Robustness Analysis

Robustness, defined here with respect to metabolic networks, is a measure of the change in the maximal flux of the objective function (the growth flux was defined as the objective) when the optimal flux through any particular metabolic reaction is changed. The robustness characteristics of the metabolic network were determined by calculating the optimal flux vector so as to maximize the growth flux (with only external flux constraints); this flux was called the *in silico* wildtype flux. Then the flux through the reaction in question was reduced from 100% to 0% of the *in silico* wildtype flux and the objective function was calculated. Additionally, the *in silico* wildtype flux was increased from the wildtype value and the upper bound on increasing the flux level was the maximal allowable flux in the reaction or the flux level for which the objective function was reduced to zero. The calculations were for a simulated aerobic batch culture in glucose minimal media.

The FBA framework was used to address the systemic effect on the metabolic network of increased and decreased (with respect to the *in silico* wildtype) metabolic flux. Herein, we quantified the robustness of the metabolic network to flux changes in the essential enzymatic reactions. The essential enzymes (for growth on glucose minimal media) were previously identified through an *in silico* analysis (24). Seven enzymatic reactions in central metabolism (Figure 1) were found to be essential: the transketolase (TKT), ribose-5-phosphate isomerase (RPI), two enzymes (GAP, PGK) in the 3-carbon stage of glycolysis (3CG), and the first three enzymes (GLT, ACN, ICD) of the TCA cycle. Below, the robustness characteristics of the metabolic network with respect to alterations of the flux levels of these essential metabolic reactions will be investigated. We will utilize phenotype phase planes (PhPPs) to define points where the optimal utilization of the metabolic network changes due to capacity constraints on the essential enzymatic reactions.

**Transketolase.** The transketolase (TKT) catalyzes an essential enzymatic reaction in the pentose phosphate pathway (PPP) (41). However, *tkt* mutant strains have

been shown to grow on glucose minimal media with low TKT residual activity (3% of wildtype) (42–43). The ability of the metabolic network to support growth with a large reduction in TKT flux was investigated *in silico* by continuously restricting the metabolic flux in the TKT reactions. As the maximum allowable flux through the TKT reactions was reduced from the *in silico* wildtype, it was determined that the ability of the metabolic network to support growth was virtually unchanged for enzymatic fluxes as low as 15% of the *in silico* wildtype (Figure 2). The response to decreased TKT metabolic flux was found to have two qualitatively different regions. The regions were identified in the PhPP (Figure 3).

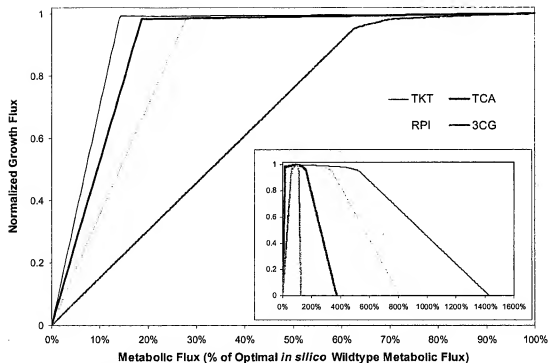
The PhPP describing the changes in the metabolic pathway utilization as a function of the TKT flux and the glucose uptake rate was calculated (Figure 3). The optimal relation between the glucose uptake rate and the TKT flux was determined from the PhPP (Figure 3). It was determined that there were two qualitatively different regions of metabolic pathway utilization for TKT fluxes lower than optimal, and these regions were defined as A and B (as shown in Figure 3). Furthermore, there were determined to be six qualitatively different regions for TKT fluxes greater than optimal, and these regions were defined as 1–6 (as shown in Figure 3). The maximal growth flux (normalized to the *in silico* wildtype) was calculated for all TKT fluxes from zero to the maximum allowable flux (the glucose uptake exchange flux was constrained to 10 mmol g-DW<sup>-1</sup> h<sup>-1</sup>) that still permits cellular growth, and the results are shown in Figure 2.

In region A (above 15% of the *in silico* wildtype enzyme flux), the optimal value of the growth flux was hardly changed, and at the demarcation between regions A and B, the growth flux was decreased to 99.2% of the *in silico* wildtype. However, to cope with the decreased TKT metabolic flux carrying capacity, shifts in the metabolic pathway utilization occurred (Figure 4). The redox potential (NADPH) requirement for biosynthetic demands was achieved by a flux redistribution that resulted in the utilization of the transhydrogenase that converted NADH (produced from an increased TCA cycle flux, Figure 4C) into NADPH. The flux diverted from the PPP (Figure 3B) resulted in increased glycolytic fluxes, in particular the pyruvate kinase and the phosphoglucosomerase fluxes (Figure 3A). In this region, the optimal growth flux was not sensitive to changes in the TKT flux. However, the optimal flux in several metabolic processes were sensitive to the TKT flux in this region (transhydrogenase, PYK, PGI, TCA cycle flux).

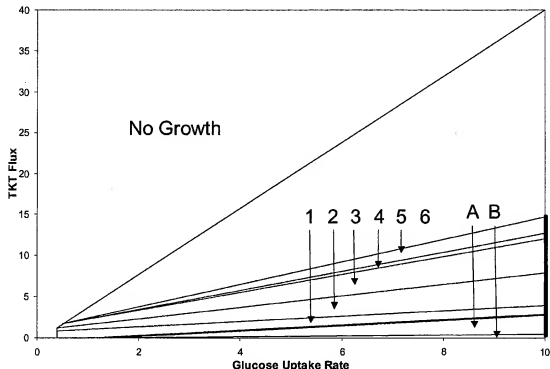
In the second region (region B) of reduced TKT flux (enzyme flux less than 15% of the *in silico* wildtype), the metabolic network was limited in the ability to produce the essential biosynthetic precursor, erythrose 4-phosphate. In this region, the optimal growth flux was sensitive to the flux level in the TKT reaction. The metabolic fluxes in this region are not shown in Figure 4 because alternate optimal solutions exist. Cellular growth is solely limited by the availability of a single biosynthetic precursor, the excess glucose can be converted to any of the metabolic byproducts with the same value of the objective function. Furthermore, the excess high-energy phosphate bonds can be eliminated in any futile cycle; thus alternate optimal solutions exist.

The effect on the metabolic network due to TKT fluxes increased beyond the optimal flux for growth was also examined. An increase in metabolic flux may result from the overexpression of the respective gene, and the robustness analysis can be used to identify the constraints on flux changes due to the integrated metabolic network.





**Figure 2.** Robustness diagrams. The effect of altered metabolic flux in the essential metabolic reactions on the normalized growth flux is illustrated. The *in silico* wildtype flux is defined as 100%. See the text for a complete discussion.

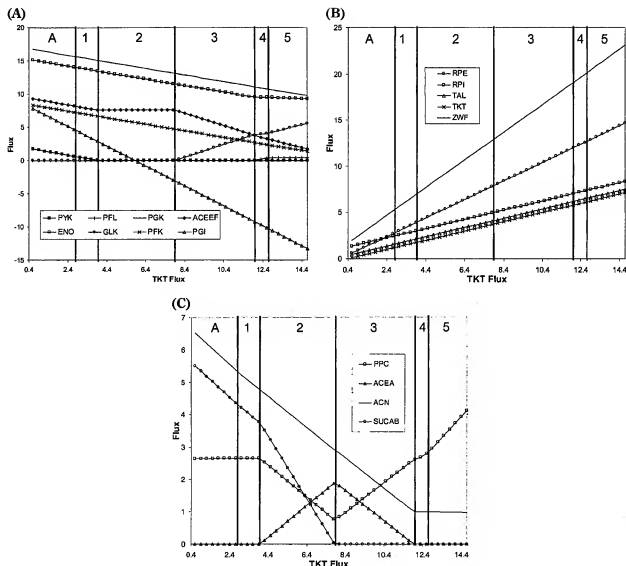


**Figure 3.** The glucose uptake rate ( $\text{mmol} \cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ )–transketolase flux (substrates converted  $\cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ ) phenotype phase plane. Exchange flux constraints were defined as discussed in the text. The regions are numbered and lettered. The numbered regions correspond to TKT fluxes that are increased relative to the optimal value. The optimal relation is the thick demarcation line. The lettered regions identify TKT flux reductions below the optimal relation. The metabolic fluxes along the thick vertical line (glucose uptake rate =  $10 \text{ mmol} \cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ ) are shown in Figure 4.

included in the analysis (such as oxygen mass transfer limitations).

**Ribose-5-Phosphate Isomerase.** The ribose-5-phosphate isomerase reaction (RPI) also catalyzes an essential reaction in the PPP (41) for growth in glucose minimal

media. Similarly to *tkt* mutants, *rpi* mutants have been shown to grow with enzymatic activity much less than that of the wildtype. For example, Skinner and Cooper have isolated a strain with RPI activity below 10% of the wildtype, and this strain was able to grow (44). As the



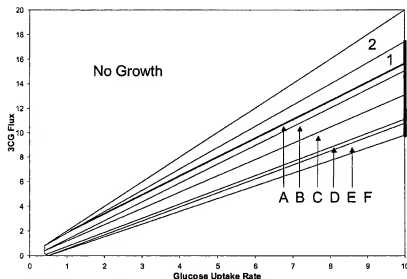
**Figure 4.** Optimal intracellular fluxes in the central metabolic pathways (substrates converted  $\cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ ) as a function of the TKT metabolic flux constraint (substrates converted  $\cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ ). The glucose uptake rate was constrained to  $10 \text{ mmol} \cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ . (A) Glycolytic fluxes. (B) Pentose phosphate pathway fluxes. (C) TCA cycle fluxes.

maximum allowable RPI flux was reduced from the *in silico* wildtype, it was determined that the ability of the metabolic network to support growth was virtually unchanged for enzymatic fluxes as low as 28% of the *in silico* wildtype (Figure 2). The RPI-glucose uptake rate PhPP was calculated to characterize the effect of altered RPI metabolic fluxes (not shown, similar to Figure 3). The qualitative effect of reduced flux in the RPI reaction from the *in silico* wildtype was investigated with FBA, and the holistic metabolic response to decreased and increased RPI fluxes was similar to the TKT results because the effect on the PPP was similar.

**3-Carbon Glycolysis.** The 3-carbon glycolytic reactions that the *in silico* analysis predicted to be essential have been shown experimentally to be required for the growth of *E. coli* on a glucose minimal media (GAP, PGK) (41). The glycolytic essential reactions were subjected to a robustness analysis to investigate the optimal systemic effect of flux alteration. The ability of the metabolic network to support growth with a reduction in 3CG flux was investigated *in silico* by continuously restricting the 3CG flux. As the allowable flux through the 3CG reac-

tions was reduced from the *in silico* wildtype, it was determined that the sensitivity of the growth flux was increased compared to the other essential reactions. When the 3CG flux was reduced below about 70% of the *in silico* wildtype, the growth flux was sensitive to the 3CG flux (Figure 2). Furthermore, the 3CG fluxes could only be increased to 110% of the *in silico* wildtype before severe limitations in the growth flux were encountered (Figure 2). We have investigated the metabolic response to 3CG flux level alterations by a phenotype phase plane analysis (Figure 5).

The PhPP describing the changes in the metabolic pathway utilization as a function of the 3CG flux and the glucose uptake rate was calculated (Figure 5). The optimal relation between the glucose uptake rate and the 3CG flux was determined from the PhPP (Figure 5). It was determined that there were six qualitatively different regions of metabolic pathway utilization for 3CG fluxes lower than optimal, and these regions were defined as A–F (as shown in Figure 5). Furthermore, there were determined to be two qualitatively different regions for 3CG fluxes greater than optimal, and these regions were



**Figure 5.** The glucose uptake rate ( $\text{mmol} \cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ )–3-carbon glycolytic flux (substrates converted  $\cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ ) phenotype phase plane. Exchange flux constraints were defined as discussed in the text. The regions are numbered and lettered. The numbered regions correspond to 3CG fluxes that are increased relative to the optimal value. The optimal relation is the thick demarcation line. The lettered regions identify 3CG flux reductions below the optimal relation. The metabolic fluxes along the thick line (glucose uptake rate =  $10 \text{ mmol} \cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ ) are shown in Figure 6.

defined as 1–2 (as shown in Figure 5). The maximal growth flux (normalized to the *in silico* wildtype) was calculated for all 3CG fluxes from zero to the maximum allowable flux (the glucose uptake exchange flux was constrained to  $10 \text{ mmol} \cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ ) that still permits cellular growth, and the results are shown in Figure 2.

The optimal relation between the glucose uptake and the 3CG flux was calculated (Figure 5). The sensitivity of other optimal fluxes in the metabolic network upon the reduction of the 3CG flux was examined (Figure 6). With 3CG flux reduction just below the optimal value, the optimal metabolic network operation was characterized by region A (Figure 5). In region A, the 3CG flux reduction led to increased PPP fluxes and the transhydrogenase was used (Figure 6B); additionally, the TCA cycle flux was reduced (Figure 6C). The reduced 3CG flux also led to the reduction of the PYK flux, which was optimally completely inactivated at the demarcation between regions A and B (Figure 6A). In region B, the glyoxylate bypass was optimally utilized and the TCA cycle fluxes were further reduced. At the demarcation between region B and C, the TCA cycle no longer operated cyclically but rather served to generate the biosynthetic precursors. In region C, glucokinase was included in the optimal flux vector, the inclusion of the glucokinase decoupled the phosphoenolpyruvate to pyruvate biochemical conversion and the uptake of glucose, thus allowing for the 3CG flux to be decreased and with little effect on the maximal growth flux. The growth flux at the demarcation between region C and D was 98% of the *in silico* wildtype (Figure 2).

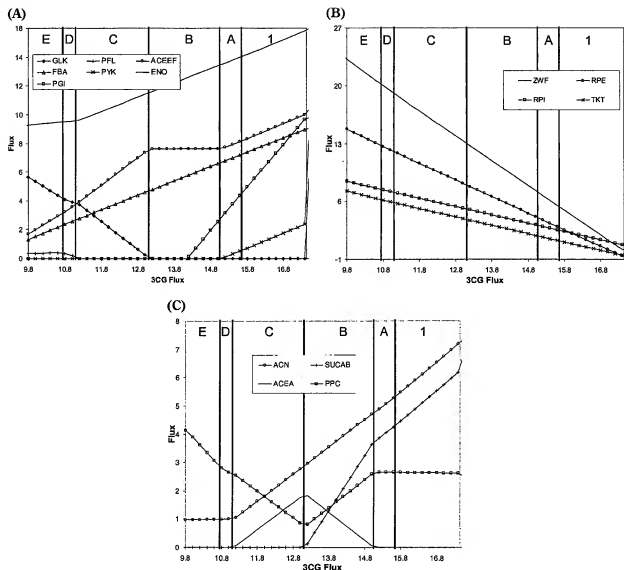
Regions D and E were very similar with respect to the optimal metabolic flux vector. In these regions, the pyruvate–formate lyase was optimally active and the glyoxylate bypass was no longer included in the optimal flux vector. Additionally, in regions D and E, the growth flux was more sensitive (compared to regions A–C) to the 3CG flux, and at the demarcation between regions E and F the maximal growth flux was 95% of the *in silico* wildtype.

In the final region of reduced 3CG flux (region F, metabolic flux less than 63% of the *in silico* wildtype), the

metabolic network was limited in the ability to produce the essential biosynthetic precursors below the block in the metabolic network. In this region, the optimal growth flux was sensitive to the 3CG flux, and the maximal growth flux linearly decreased to zero as the 3CG flux was reduced to zero from the region E,F boundary. The metabolic fluxes in this region are not shown in Figure 6 because alternate optimal solutions exist. Cellular growth is limited by the availability of the biosynthetic precursors after the metabolic blockage, and the diversion of the flux from glycolysis to the PPP resulted in excess high-energy phosphate bonds and redox potential. The growth flux in region F is dependent upon increased oxygen availability to eliminate the excess redox potential. An additional constraint was imposed on the metabolic network, i.e., the oxygen uptake was constrained below a physiologically realistic value of  $20 \text{ mmol} \cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ , and the feasible set did not contain a growth flux for 3CG fluxes below about 40% of the *in silico* wildtype (not shown). Thus, the partial inhibition of the 3CG fluxes can theoretically prevent the growth of *E. coli*; however, growth can be maintained with reduced glucose uptake rates.

The holistic effect of increased 3CG flux on the metabolic networks capability to support cellular growth was assessed with FBA. The optimal growth flux as a function of the flux in the 3CG flux was calculated (Figure 2). The metabolic flux was continuously increased *in silico* from the *in silico* wildtype value to the maximum flux that permits growth, and two qualitatively different metabolic flux vectors (numbered 1 and 2 in Figure 5) were observed.

In region 1 (Figure 5), the optimal metabolic flux vector was characterized by a decreased pentose phosphate pathway (PPP), an active transhydrogenase reaction, an increased PYK flux, and an increased TCA cycle flux. However, region 1 only extends to a 3CG flux of 110% of the *in silico* wildtype (with a glucose uptake of  $10 \text{ mmol} \cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ ), and region 2 of Figure 5 was characterized by alternate optimal flux distributions. The metabolic network was limited in the ability to produce the essential biosynthetic precursors before the effected reac-



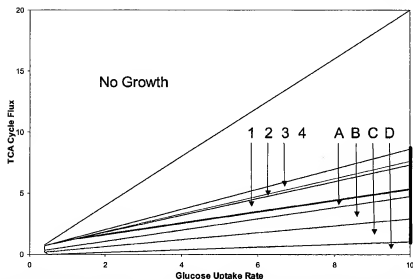
**Figure 6.** Optimal intracellular fluxes (substrates converted  $\cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ ) in the central metabolic pathways as a function of the 3CG metabolic flux constraint (substrates converted  $\cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ ). The glucose uptake rate was constrained to  $10 \text{ mmol} \cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ . (A) Glycolytic fluxes. (B) Pentose phosphate pathway fluxes. (C) TCA cycle fluxes.

tions of 3-carbon glycolysis. In this region, the optimal glucose flux was sensitive to the 3CG flux. The metabolic fluxes in this region are not shown in Figure 6 because alternate optimal solutions exist. The excess pyruvate produced by the elevated 3CG flux can be converted to any of the metabolic byproducts with the same value of the objective function. Furthermore, the excess high-energy phosphate bonds can be eliminated in any futile cycle, and thus alternate optimal solutions exist.

**TCA Cycle.** The initial three fluxes of the TCA cycle were determined to be essential. The deletion of any of these enzymatic activities resulted in a glutamate requirement. This requirement has been shown experimentally (45). The *in silico* robustness analysis was performed to assess the effect of decreased (and increased) flux entering the TCA cycle (Figures 2, 7, and 8). The ability of the metabolic network to support growth with a reduction in the TCA cycle flux was investigated *in silico* by continuously restricting the citrate synthase flux, which we will refer to as the TCA cycle flux. As the TCA cycle flux constraint was reduced from the *in silico* wildtype, it was determined that the ability of the

metabolic network to support growth was not sensitive to the TCA cycle flux above 18% of the *in silico* wildtype. Furthermore, the TCA cycle flux could be increased to about 160% of the *in silico* wildtype before severe limitations in the growth flux were encountered (Figure 2). We have investigated the metabolic response to TCA cycle flux level alterations by a phenotype phase plane analysis (Figure 7).

The PhPP describing the changes in the metabolic pathway utilization as a function of the TCA cycle flux and the glucose uptake rate was calculated (Figure 7). The optimal relation between the glucose uptake rate and the TCA cycle flux was determined from the PhPP (Figure 7). It was determined that there were four qualitatively different regions of optimal metabolic pathway utilization for TCA cycle fluxes lower than optimal, and these regions were defined as A–D (as shown in Figure 7). Furthermore, there were determined to be four qualitatively different regions for TCA cycle fluxes greater than optimal, and these regions were defined as 1–4 (as shown in Figure 7). The maximal growth flux (normalized to the *in silico* wildtype) was calculated for all TCA cycle



**Figure 7.** The glucose uptake rate ( $\text{mmol} \cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ )–TCA cycle flux (substrates converted  $\cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ ) phenotype phase plane. Exchange flux constraints were defined as discussed in the text. The regions are numbered and lettered. The numbered regions correspond to TCA cycle fluxes that are increased relative to the optimal value. The optimal relation is the thick demarcation line. The lettered regions identify TCA cycle flux reductions below the optimal relation. The metabolic fluxes along the thick line (glucose uptake rate =  $10 \text{ mmol} \cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ ) are shown in Figure 8.

fluxes from zero to the maximum allowable flux (The glucose uptake exchange flux was constrained to  $10 \text{ mmol} \cdot \text{g-DW}^{-1} \cdot \text{h}^{-1}$ ) that still permits cellular growth, and the results are shown in Figure 2.

The optimal relation between the glucose uptake and the TCA cycle flux was calculated, and the sensitivity of the optimal fluxes in central metabolism to the TCA cycle flux was examined (Figure 8). Region A defines the optimal set of metabolic reactions that are utilized with a reduction of the TCA cycle flux below the optimal value (Figure 7). In region A, the reduction of the TCA cycle flux led to increased PPP fluxes and the transhydrogenase was used; additionally, the glycolytic flux was decreased. The reduced TCA cycle flux led to the reduction of the PYK flux, which was optimally completely inactivated at the demarcation between regions A and B. In region B, the glyoxylate bypass was optimally utilized and the glycolytic fluxes were further reduced. At the demarcation between region B and C, the TCA cycle fluxes were reduced to the point that the TCA cycle no longer operated cyclically but rather served to generate the biosynthetic precursors. In region C, glucokinase was included in the optimal flux vector, and the inclusion of the glucokinase decoupled the phosphoenolpyruvate to pyruvate biochemical conversion and the uptake of glucose. The growth flux at the demarcation between region C and D was 98% of the *in silico* wildtype (Figure 2) and the TCA cycle flux was 18% of the *in silico* wildtype TCA cycle flux.

In the final region of reduced TCA cycle fluxes (metabolic flux less than 18% of the *in silico* wildtype), the metabolic network was limited in the ability to produce  $\alpha$ -ketoglutarate, an essential biosynthetic precursor. In this region, the optimal growth flux was sensitive to the TCA cycle flux, and the characteristic behavior was similar to region F of the 3CG-glucose uptake rate PhPP (Figure 5) that was discussed above.

The holistic effect of increased TCA cycle flux on the metabolic networks' capability to support cellular growth was assessed with FBA. The optimal growth flux as a function of the flux in the TCA cycle flux was calculated (Figure 2). The metabolic flux was continuously increased

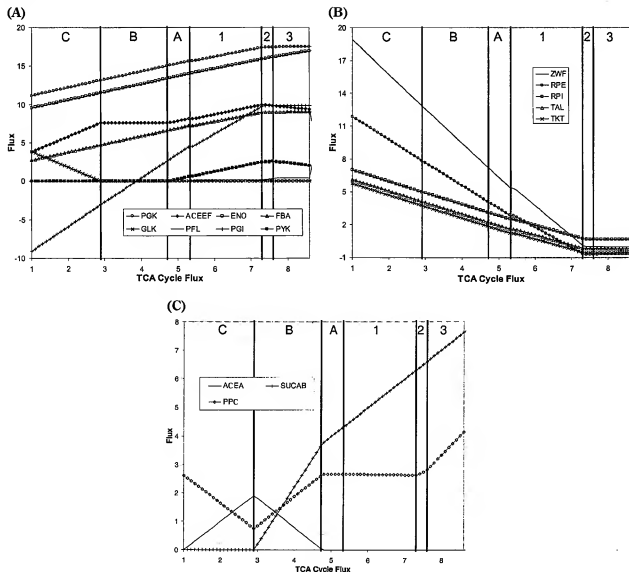
*in silico* from the *in silico* wildtype value to the maximum flux that permits growth, and four qualitatively different metabolic flux vectors (numbered 1–4 in Figure 7) were observed.

The region 1 of Figure 7 the optimal metabolic flux vector was characterized by an active transhydrogenase reaction, an increased PYK flux, an increased glycolytic flux, and a decreased PPP (which was optimally inactivated at the demarcation between regions 1 and 2). Regions 2 and 3 are very similar with respect to the set of metabolic reactions that are optimally utilized, and in these regions, the PPP is optimally inactivated. At the demarcation between regions 3 and 4, the maximal growth flux was about 95% of the *in silico* wildtype and the TCA cycle flux was increased to approximately 160% of the *in silico* wildtype. With TCA cycle flux increases beyond region 3, region 4 is encountered (Figure 7). Region 4 was characterized by alternate optimal flux distributions. The metabolic network was limited in the ability to produce the essential glycolytic and PPP biosynthetic precursors. In this region, the optimal growth flux was sensitive to the TCA cycle flux, and the metabolic fluxes in this region are not shown in Figure 8 because alternate optimal solutions exist.

## Discussion

We have illustrated, with the complete *E. coli* metabolic network, how optimal metabolic phenotypes (flux vectors) and shifts in metabolic behavior can be analyzed and interpreted *in silico*. From the fundamental physicochemical constraints on the metabolic network, the feasible set that identifies the capabilities of the metabolic network was identified. Subsequently, a linear optimization routine was utilized to search the feasible set for a flux vector that maximizes a given objective function. Given the complexity associated with developing complete dynamic modeling of cellular processes, the constraining approach, as discussed herein, is a particularly useful alternative approach to metabolic systems analysis. The results presented herein are of fundamental interest for several reasons. First, the ability to define essential genes under various conditions will have many





**Figure 8.** Optimal intracellular fluxes (substrates converted  $\cdot$  g-DW $^{-1} \cdot$  h $^{-1}$ ) in the central metabolic pathways as a function of the TCA cycle metabolic flux constraint (substrates converted  $\cdot$  g-DW $^{-1} \cdot$  h $^{-1}$ ). The glucose uptake rate was constrained to 10 mmol g-DW $^{-1}$  h $^{-1}$ . (A) Glycolytic fluxes. (B) Pentose phosphate pathway fluxes. (C) TCA cycle fluxes.

practical applications. Second, the results presented discussed the sensitivity of the objective function to specific fluxes in the metabolic network. Finally, the results demonstrate the potential capabilities of *in silico* analysis of cellular systems. Understanding the relation between individual fluxes and the holistic function of the metabolic network is essential to successfully metabolic engineering a living system, and FBA provides a methodology that can be used to direct the metabolic engineer.

First, we should address the assumptions associated with the utilization of linear optimization to identify the optimal flux vector. The metabolic constraining formalism that we have discussed is based on the fundamental physicochemical constraints that all cells must abide to. Within the set of constraints, the cell will choose a flux vector for which to operate. We have attempted to find the same flux vector by employing linear optimization, and the assumption is that the cell has evolved the regulatory mechanisms to find the optimal solution within the physicochemical constraints to maximize its survival. We have mathematically represented survival as cellular growth. On the basis of comparisons of the linear

optimization results and experimental data, the assumption appears to be valid under the tested conditions (27). However, currently, the number of situations for which the validity of the assumption has been addressed is limited. Therefore, further experimental validation is in order. It should be noted that, even if the assumption of optimal growth proves correct for wildtype strains, it is not clear if we should expect that an engineered strain will behave in an optimal manner. Therefore, the optimization results may only provide an upper bound on the expected behavior of engineered strains.

The identification of the essential gene products in a metabolic network is of fundamental interest (46, 47). The metabolic constraining formalism that was described here provides an efficient method to study the consequences of alterations in the genotype and to gain insight into the genotype-phenotype relation. The study of the removal of individual metabolic enzymes in the central metabolic pathways demonstrated fundamental redundancy properties of the *E. coli* metabolic genotype and the existence of relatively few critical gene products. Seven metabolic reactions were determined to be essential.

ential in the metabolic network. Of these seven reactions, one set of three forms a linear reaction series and another set of two forms another linear reaction series; therefore, the effect of deleting any of the reactions in the linear set is equivalent. Thus, there were basically four different metabolic fluxes that were essential: two in the PPP, the 3-carbon stage of glycolysis, and the first three reactions of the TCA cycle. It was shown that there are regions where the ability of the metabolic network to support growth was not affected by the altered flux levels in the essential enzymatic reactions. This metabolic robustness was due to the capability of the metabolic network to shift the production of redox potential and high-energy phosphate bonds. However, the point where the metabolic network was unable to support growth was quantitatively derived *in silico*. At this point, the metabolic network was limited in the ability to produce biosynthetic precursors. In the biosynthetic precursor limited region, the availability of redox potential and high-energy phosphate bonds did not limit the growth capability of the cell.

The robustness analysis presented herein is essentially a sensitivity analysis. The sensitivity of the objective function to changes in the optimal flux vector was addressed. A constraint was added to the metabolic network, thus altering a single metabolic flux, and a new optimal flux vector was calculated. The relation between the additional constraint and the objective function was examined to investigate the robustness in the system with respect to the essential enzymatic reactions. Furthermore, the relation between the additional constraint and optimal flux vector was also examined. The results of the sensitivity analysis can be used to interpret the optimal metabolic fluxes and their relation to *in vivo* metabolic fluxes. For example, metabolic fluxes for which the objective function is highly sensitive are likely to obtain *in vivo* values that are consistent with the optimal values. Furthermore, metabolic fluxes to which the objective is not sensitive may be able to take on flux values in a large range with very little effect on the optimal solution. Additionally, other fluxes in the metabolic network that are sensitive to a given metabolic flux may provide an indication of the accuracy of the optimal predictions. For example, the TCA cycle flux was reduced to approximately 18% of the *in silico* wildtype with little effect on the objective function; however, the PPP fluxes were three times the *in silico* wildtype, fluxes much higher than experimental data indicates.

Understanding the metabolic fluxes and their control is essential to the ability to "design" metabolic networks for the production commodity chemicals (i.e., antibiotics, vitamins, amino acids, etc). Using flux balance analysis the complete range of metabolic phenotypes can be examined under defined environmental and genetic conditions through the use of phenotypic phase planes (PhPP). For the *E. coli* network PhPPs were generated for growth on glucose minimal media spanning the uptake rate of the carbon source and an intracellular metabolic flux. Therefore, the PhPP formalism has provided an efficient methodology for examining the consequence of altered fluxes within the cell. Therefore, bioinformatically based models will undoubtedly have a major impact on the development of metabolic engineering (4, 5). Herein, we have investigated the effect of altered metabolic flux levels on the maximal growth flux, thus quantifying the relation between altered flux levels and optimal cellular growth. Furthermore, by examining the entire set of constraints on the metabolic network, constraints on metabolic flux alterations can be identified.

FBA incorporates no information on enzyme kinetics or gene regulation, thus limiting insight into dynamic responses. From flux balance analysis it is possible to realize some of the fundamental constraints that metabolic systems are faced with and define the feasible set that contains all admissible steady-state flux vectors. As *in vivo* reaction dynamics is further understood, the ability to predict dynamic responses of metabolic networks to environmental and genetic perturbations using dynamic modeling approaches will become more feasible. In general, regulatory schemes and reaction dynamics will serve to further constrain metabolic behavior to operate in confined subspaces of the feasible set. Identifying these regions from both the theoretical and experimental side will be a challenge for the future.

A number of experimental technologies have now made the holistic study of biological systems feasible. The ability to assimilate DNA chip-based and protein expression technologies providing genome-scale information with computational methods for metabolic network analysis will become important in advancing the study of metabolic physiology and the practice of metabolic engineering. Currently the interpretation of high-throughput experimental information on systemic behavior is limited by a lack of analysis capabilities. Can systems-based quantitative *in silico* approaches such as flux balance be used to assist in understanding this flood of data? This question will need to be answered as interest builds in the genomics community for quantitative systems analysis.

The analysis of the metabolic phenotype-genotype relation using the bioinformatically based *in silico* metabolic genotype of *E. coli* can serve as a basis for the construction of parallel *in silico* representations of other single-cell organisms. Thus, the results presented are particularly relevant with the current emphasis on genome sequencing. Utilizing the techniques described herein, information can be gained regarding the metabolic physiology of a cell with relatively little experimental biochemical information on the cell of interest. However, this analysis should be considered a single step toward the integrative analysis of bioinformatic databases to predict and understand cellular function based on the underlying genetic content. Continued prediction and experimental verification will be an integral part of the further development of *in silico* strains and their use to represent their *in vivo* counterparts.

### Acknowledgment

Portions of this work are supported by grants from the National Institute of Health (57089-01A1) and the National Science Foundation (9873384).

### References and Notes

- Bhalla, U. S.; Iyengar, R. Emergent properties of networks of biological signaling pathways. *Science* **1999**, *283*, 381–387.
- Weng, G.; Bhalla, U. S.; Iyengar, R. Complexity in biological signaling systems. *Science* **1999**, *284*, 92–96.
- Kanehisa, M. Databases of biological information. *Trends Guide Bioinf.* **1998**, 24–26.
- Palsson, B. O. What lies beyond bioinformatics? *Nat. Biotechnol.* **1997**, *15*, 3–4.
- Edwards, J. S.; Palsson, B. O. How will bioinformatics influence metabolic engineering? *Biotechnol. Bioeng.* **1998**, *58*, 162–169.
- Goodwin, B. C. *Oscillatory Organization in Cells, a Dynamic Theory of Cellular Control Processes*; Academic Press: New York, 1963.

- (7) Tyson, J. J.; Othmer, H. G. The dynamics of feedback control circuits in biochemical pathways. *Prog. Theor. Biol.* **1978**, *5*, 1–62.
- (8) Savageau, M. The behaviour of intact biochemical control systems. *Curr. Top. Cell Regul.* **1972**, *6*, 64–130.
- (9) Hess, B.; Boiteux, A. Oscillatory organization in cells, a dynamic theory of cellular control processes. *Hoppe-Seyler's Z. Physiol. Chem.* **1968**, *349*, 1567–1574.
- (10) Savageau, M. A. Biochemical systems analysis. II. The steady-state solutions for an  $n$ -pool system using a power-law approximation. *J. Theor. Biol.* **1969**, *25*, 370–379.
- (11) Heinrich, R.; Rapoport, S. M.; Rapoport, T. A. Metabolic regulation and mathematical models. *Prog. Biophys. Mol. Biol.* **1977**, *32*, 1–82.
- (12) Kacser, H.; Burns, J. A. The control of flux. *Symp. Soc. Exp. Biol.* **1973**, *27*, 65–104.
- (13) Fell, D. *Understanding the Control of Metabolism*; London: Portland Press: 1996.
- (14) Lee, I. D.; Pálsson, B. O. A Macintosh software package for simulation of human red blood cell metabolism. *Comput. Methods Prog. Biomed.* **1992**, *38*, 195–226.
- (15) Joshi, A.; Pálsson, B. O. Metabolic dynamics in the human red cell. Part I: A comprehensive kinetic model. *J. Theor. Biol.* **1999**, *141*, 515–528.
- (16) Mulquhney, P. J.; Kuchel, P. W. Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: computer simulation and metabolic control analysis. *Biochem. J.* **1999**, *342*, Pt 3, 597–604.
- (17) Pálsson, B. O.; Joshi, A.; Öztürk, S. S. Reducing complexity in metabolic networks: making metabolic meshes manageable. *Fed. Proc.* **1987**, *46*, 2485–2489.
- (18) Bailey, J. E. Mathematical modeling and analysis in biochemical engineering: Past accomplishments and future opportunities. *Biotechnol. Prog.* **1998**, *14*, 8–20.
- (19) Bonarius, H. P. J.; Schmid, G.; Tramper, J. Flux analysis of underdetermined metabolic networks: The quest for the missing constraints. *Trends Biotechnol.* **1997**, *15*, 308–314.
- (20) Edwards, J. S.; Ramakrishna, R.; Schilling, C. H.; Pálsson, B. O. Metabolic Flux Balance Analysis. In *Metabolic Engineering*; Lee, S. Y., Papoutsakis, E. T., Eds.; Marcel Dekker: New York, 1999; pp 13–57.
- (21) Edwards, J. S.; Pálsson, B. O. Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. *J. Biol. Chem.* **1999**, *274*, 17410–17416.
- (22) Varma, A.; Pálsson, B. O. Metabolic Flux Balancing: Basic concepts, Scientific and Practical Use. *Bio/Technology* **1994**, *12*, 994–998.
- (23) Sauer, U.; Cameron, D. C.; Bailey, J. E. Metabolic capacity of *Bacillus subtilis* for the production of purine nucleosides, riboflavin, and folic acid. *Biotechnol. Bioeng.* **1998**, *59*, 227–238.
- (24) Edwards, J. S.; Pálsson, B. O. The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5528–5533.
- (25) Pálsson, B. O.; Lightfoot, E. N. Mathematical modelling of dynamics and control in metabolic networks. Part I. On Michaelis–Menten kinetics. *J. Theor. Biol.* **1984**, *111*, 273–302.
- (26) Clarke, B. L. Stoichiometric network analysis. *Cell Biophys.* **1988**, *12*, 237–253.
- (27) Varma, A.; Pálsson, B. O. Stoichiometric flux balance models quantitatively predict growth and metabolic byproduct secretion in wildtype *Escherichia coli* W3110. *Appl. Environ. Microbiol.* **1994**, *60*, 3724–3731.
- (28) Walsh, K.; Koshland, D. Determination of flux through the branch point of two metabolic cycles. *J. Biol. Chem.* **1984**, *259*, 9646–9654.
- (29) Karp, P. D.; Riley, M.; Saier, M.; Paulsen, I. T.; Paley, S. M.; Pellegrini-Toole, A. The EcoCyc and MetaCyc databases. *Nucleic Acids Res.* **2000**, *28*, 56–9.
- (30) Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **1999**, *27*, 29–34.
- (31) Selkov, E. Jr.; Grechkin, Y.; Mikhailova, N.; Selkov, E. MPW: the Metabolic Pathways Database. *Nucleic Acids Res.* **1998**, *26*, 43–45.
- (32) Neidhardt, F. C.; Umbarger, H. E. Chemical Composition of *Escherichia coli*. In *Escherichia coli and Salmonella: cellular and molecular biology*, 2nd ed.; Neidhardt, F. C., Ed.; ASM Press: Washington, DC, 1996; Vol. 1, pp 13–16.
- (33) Pramanik, J.; Keasling, J. D. Stoichiometric model of *Escherichia coli* metabolism: Incorporation of growth-rate dependent biomass composition and mechanistic energy requirements. *Biotechnol. Bioeng.* **1997**, *56*, 398–421.
- (34) Varma, A.; Pálsson, B. O. Parametric sensitivity of stoichiometric flux balance models applied to wildtype *Escherichia coli* metabolism. *Biotechnol. Bioeng.* **1995**, *45*, 69–79.
- (35) Varma, A.; Pálsson, B. O. Predictions for Oxygen Supply Control to Enhance Population Stability of Engineered Production Strains. *Biotechnol. Bioeng.* **1994**, *43*, 275–285.
- (36) Schilling, C. H.; Pálsson, B. O. The underlying pathway structure of biochemical reaction networks. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4193–4198.
- (37) Schilling, C. H.; Edwards, J. S.; Pálsson, B. O. Towards metabolic phenomics: Analysis of genomic data using flux balances. *Biotechnol. Prog.* **1999**, *15*, 288–95.
- (38) Schilling, C. H.; Schuster, S.; Pálsson, B. O.; Heinrich, R. Metabolic pathway analysis: basic concepts and scientific applications in the post-genomic era. *Biotechnol. Prog.* **1999**, *15*, 296–303.
- (39) Edwards, J. S.; Ramakrishna, R.; Pálsson, B. O. *Characterizing phenotypic plasticity: A phase plane analysis*, submitted for publication.
- (40) Varma, A.; Pálsson, B. O. Metabolic capabilities of *Escherichia coli*: I. Synthesis of biosynthetic precursors and cofactors. *J. Theor. Biol.* **1993**, *165*, 477–502.
- (41) Fraenkel, D. G. Glycolysis. In *Escherichia coli and Salmonella*; Neidhardt, F. C., Ed.; ASM Press: Washington, DC, 1996; Vol. 1, pp 189–198.
- (42) Josephson, B. L.; Fraenkel, D. G. Sugar metabolism in transketolase mutants of *Escherichia coli*. *J. Bacteriol.* **1974**, *118*, 1082–1089.
- (43) Josephson, B. L.; Fraenkel, D. G. Transketolase mutants of *Escherichia coli*. *J. Bacteriol.* **1969**, *100*, 1289–1295.
- (44) Skinner, A. J.; Cooper, R. A. Genetic studies on ribose 5-phosphate isomerase mutants of *Escherichia coli* K-12. *J. Bacteriol.* **1974**, *118*, 1183–1185.
- (45) Cronan, J. E., Jr.; Laporte, D. Tricarboxylate acid cycle and glyoxylate bypass. In *Escherichia coli and Salmonella*; Neidhardt, F. C., Ed.; ASM Press: Washington, DC, 1996; Vol. 1, pp 189–198.
- (46) Arigoni, F.; Talabot, F.; Peitsch, M.; Edgerton, M. D.; Meldrum, E.; Allet, E.; Fish, R.; Jamotte, T.; Curchod, M. L.; Leferer, H. A genome-based approach for the identification of essential bacterial genes. *Nat. Biotechnol.* **1998**, *16*, 851–856.
- (47) Akterley, B. J.; Rubin, E. J.; Camilli, A.; Lampe, D. J.; Robertson, H. M.; Mekalanos, J. J. Systematic identification of essential genes by *in vitro* mariner mutagenesis. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8927–8932.

Accepted for publication June 16, 2000.

BP0000712

# The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*

Adam M Feist<sup>1</sup> & Bernhard Ø Palsson<sup>1,2</sup>

The number and scope of methods developed to interrogate and use metabolic network reconstructions has significantly expanded over the past 15 years. In particular, *Escherichia coli* metabolic network reconstruction has reached the genome scale and been utilized to address a broad spectrum of basic and practical applications in five main categories: metabolic engineering, model-directed discovery, interpretations of phenotypic screens, analysis of network properties and studies of evolutionary processes. Spurred on by these accomplishments, the field is expected to move forward and further broaden the scope and content of network reconstructions, develop new and novel *in silico* analysis tools, and expand in adaptation to uses of proximal and distal causation in biology. Taken together, these efforts will solidify a mechanistic genotype-phenotype relationship for microbial metabolism.

The availability of reconstructed metabolic networks for microorganisms has increased rapidly in recent years, and a growing number of research groups are reconstructing metabolic networks for organisms of interest<sup>1</sup>. A network reconstruction represents a highly curated set of primary biological information for a particular organism and thus can be considered a biochemically, genetically and genomically structured (BiGG) database<sup>1,2</sup>. A curated BiGG database (*de facto* a knowledge base) can be converted into a mathematical format (that is, an *in silico* model), and used to computationally assess phenotypic properties using a variety of computational methods<sup>2,3</sup>. Genome-scale reconstructions are thus a key step in quantifying the genotype-phenotype relationship and can be used to 'bring genomes to life'<sup>4</sup>.

The purpose of this review is to summarize and classify applications using the *E. coli* reconstruction to answer a broad spectrum of biological questions. By doing so, we provide both an up-to-date overview of the applications of constraint-based analysis and a guide to similar applications for the growing number of organisms for which genome-scale reconstructions are becoming available.

## Model formulation and the *E. coli* metabolic reconstruction

The four key steps in the formulation and use of genome-scale models are illustrated in Figure 1. Foundational to the process is the generation of global, or genome-scale, 'omics data'. Omics data, along with legacy information (that is, the 'bibliome') and small-scale detailed experiments, can be used to define the interactions among the biological components that are used to reconstruct organism-specific networks<sup>1</sup>. Network reconstruction is also an iterative, ongoing process that continually integrates data in a formal fashion as they become available<sup>5</sup>. As a result, a current and well-curated genome-scale network reconstruction is a common denominator

for those studying systems biology of an organism; for an in-depth review of the bottom-up reconstruction process, see ref. 2.

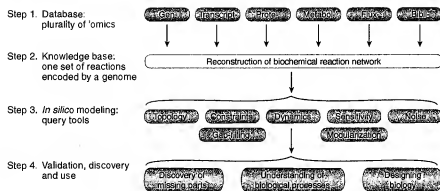
The arrow from step 2 to step 3 in Figure 1 involves a somewhat subtle, but critical, transition. With the definition of systems boundaries and other details, a network reconstruction can be converted into a mathematical format that can be computationally interrogated and subsequently used for experimental design<sup>3</sup>. Thus, a network reconstruction is converted into a genome-scale model (GEM)<sup>3</sup>. This arrow represents a bridge between the realms of high-throughput data/bioinformatics on the one hand and systems science on the other. A network reconstruction (or a BiGG knowledgebase) is, in principle, accessible to all and significant strides have been made to make computation with GEMs more readily accessible<sup>6-11</sup>. This availability of both genome-scale reconstruction and GEMs has unleashed creativity in research groups around the world and resulted in the series of studies reviewed below.

The 18-year history of reconstruction of the *E. coli* metabolic network (summarized in Fig. 2) has culminated in a network containing a total number of 1,260 open reading frame (ORF) metabolic functions<sup>12-19</sup>. This reconstruction represents 48% of the experimentally determined ORF functions in the *E. coli* genome (Table 1). It should be noted that the functions of 92% of the 1,260 gene products have been experimentally verified. Reconstruction of the *E. coli* network has thus approached an exhaustion of known metabolic gene functions and it is now being used in a prospective fashion to discover new metabolic capabilities (see below). The reconstruction of the *E. coli* metabolic network represents the best-developed genome-scale network to date and it has proven to be a platform for a variety of computational analyses. It should be noted that although there are different *E. coli* GEMs, there is only one unique network for *E. coli* and each successive iteration strives to best represent this content (we use 'the *E. coli* GEM' to refer to any of these iterations). Three successive *E. coli* GEMs from our group<sup>17-19</sup> have been used as the basis for over 60 detailed studies reviewed below.

A growing number of research groups use the *E. coli* GEM for prediction, interpreting and understanding *E. coli* phenotypic states and function.

<sup>1</sup>Department of Bioengineering, University of California San Diego, 9500 Gilman Drive, La Jolla, California 92093-0412, USA. <sup>2</sup>Center for Systems Biology, University of Iceland, Vatnsmyrarveg 16, IS-101, Reykjavik, Iceland. Correspondence should be addressed to B.Ø.P. (palsson@ucsd.edu).

Published online 6 June 2008; doi:10.1038/nbt1401



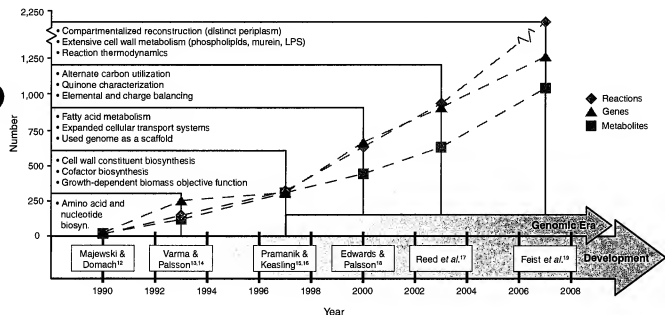
**Figure 1** Formulation and use of GEMs as a four-step process. Step 1, the process is based on a variety of high-throughput data sets (that is, omics data) and a comprehensive assessment of the literature (that is, bibliomic data). Step 2, all of the data types are used to reconstruct the list of biochemical transformations that make up a network, as well as their genetic basis<sup>1</sup>. In principle, the network is unique. Step 3, the data contained in the reconstruction can be formally represented (that is, in the form of matrices and logical statements) that can be mathematically characterized by a variety of methods. Step 4, the computational model enables a broad spectrum of applications, as reviewed in this article. Figure adapted from ref. 2.

In addition, the reconstruction itself has been used as a context for the interpretation of large amounts of experimental data. Applications of the *E. coli* GEM range from pragmatic to theoretical studies, and can be classified into five general categories (Fig. 3): first, metabolic engineering<sup>20–30</sup>; second, biological discovery<sup>31–37</sup>; third, assessment of phenotypic behavior<sup>19,38–43</sup>; fourth, biological network analysis<sup>44–79</sup>; and fifth, studies of bacterial evolution<sup>80–82</sup>. The *in silico* methods used to probe the *E. coli* GEM in each study are summarized in Figure 4. These methods perform an assessment of the solution spaces associated with the mathematical representation of a reconstruction<sup>2</sup>; they are categorized as either unbiased or biased<sup>3</sup>. The latter category relies on an observer bias that is stated through an objective function (that is now beginning to be experimentally examined<sup>83</sup>) that has been utilized in most of the studies reviewed here for the general application of flux balance analysis (FBA)<sup>84–86</sup>. Each category

of application is now detailed, with emphasis on the first three that have the greatest practical utility.

### GEMs and metabolic engineering

Through the application of computational methods that incorporate linear, mixed integer linear and nonlinear programming, it has been demonstrated that model-directed strain design can lead to increased metabolite production<sup>20–30</sup>. In these studies, the *E. coli* GEM is principally used to analyze the metabolite production potential of *E. coli* and identify metabolic interventions needed to produce the metabolite of interest. Thus, *E. coli* strains have been systematically designed through *in silico* analysis to overproduce target metabolites, such as lycopene<sup>23,24</sup>, lactic acid (our group<sup>25</sup>), ethanol<sup>26</sup>, succinic acid<sup>27,28</sup>, L-valine<sup>29</sup>, L-threonine<sup>30</sup>, additional amino acids<sup>21</sup> as well as diverse products from hydrogen to vanillin<sup>22</sup>.



**Figure 2** The iterative reconstruction and history of the *E. coli* metabolic network. Six milestone efforts are shown that contributed to the reconstruction of the *E. coli* metabolic network. For each of the six reconstructions<sup>12–19</sup>, the number of included reactions (blue diamonds), genes (green triangles) and metabolites (purple squares) are displayed. Also listed are noteworthy expansions that each successive reconstruction provided over previous efforts. For example, Varma & Palsson<sup>13,14</sup> included amino acid and nucleotide biosynthesis pathways in addition to the content that Majewski & Domach<sup>12</sup> characterized. The start of the genomic era<sup>92</sup> (1997) marked a significant increase in included reconstruction components for each successive iteration. The reaction, gene and metabolite values for pre-genomic-era reconstructions were estimated from the content outlined in each publication and in some cases, encoding genes for reactions were unclear.

Select exemplary metabolic engineering applications are described in more detail below.

To increase the production of an already high producing strain, a systematic computational search was developed<sup>24</sup> to explore the *E. coli* metabolic network and report gene deletions that diverted metabolic flux toward the desired product lycopene. This process resulted in a knockout strain, that when constructed, showed a twofold increase in the production of lycopene over the parental strain. In this analysis, the computational algorithm MOMA (minimization of metabolic adjustment)<sup>41</sup> and the *i*JE660 (ref. 18) *E. coli* GEM were used to sequentially examine additive genetic deletions that would improve lycopene production while maintaining cell viability. Strain designs were constructed through genetic manipulations using the predicted modifications and this computational approach yielded a twofold increase in production rate over a previously engineered overproducing strain and an 8.5-fold increase over wild-type production harboring only a lycopene biosynthesis plasmid<sup>24</sup>. Strain performance was evaluated by monitoring lycopene production through enzymatic assays and mutant growth rates. In addition, when the strain designs identified computationally were compared with mixed combinatorial transposon mutagenesis, the maximum observed production could be designed solely using the systematic GEM-aided computational method<sup>23,24</sup>. Furthermore, a deleterious effect was observed when targets identified in individual computational designs were combined in an attempt to achieve an overall more desirable phenotype. Thus, the overall systematic effects from individual designs were not additive and needed to be interpreted in the context of the entire network.

Two studies producing the amino acids L-valine<sup>29</sup> and L-threonine<sup>30</sup> have demonstrated the broad usage of GEM-aided computation for strain design. To generate an L-threonine production strain, GEM-aided modeling was employed in three different areas to increase production to industrial titers<sup>30</sup>. In one instance, *in silico* parametric sensitivity analysis that compared reaction activity to L-threonine production rate was used to identify the optimal activity of a key enzymatic reaction associated with maximal L-threonine production. The optimal activity prediction was subsequently used to tune the overexpression of the gene encoding the enzyme involved in this reaction through comparison to base-line activity and the result was a production increase. This method proved to be vital to the success of this strain, as a previous transcription-profiling guided attempt at overexpression resulted in an undesirable surplus of activity and was detrimental to L-threonine production. For the same strain, a GEM-aided flux analysis in conjunction with mRNA expression data levels also was used as a guide for elimination of negative regulation of a gene encoding an enzyme involved in a reaction that channeled flux toward the final product. The third use of the GEM for the design of this strain occurred when an unwanted byproduct was observed in the culture medium and computation was used to divert the flux from this byproduct to L-threonine<sup>30</sup> through overexpression of another key gene-encoded activity.

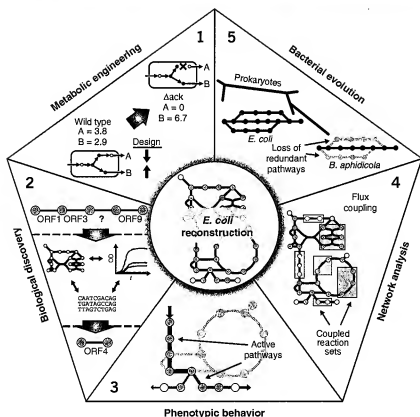
**Table 1 Properties of the most current *E. coli* metabolic reconstruction<sup>19</sup>**

Included genes	1,260	(28%) <sup>a</sup>
Experimentally based function	1,161	(92%) <sup>a</sup>
Computationally predicted function	99	(8%)
Unique functional proteins	1,148	
Multigene complexes	167	
Genes involved in complexes	415	
Instances of isozymes <sup>b</sup>	346	
Reactions	2,077	
Metabolic reactions	1,387	
Unique metabolic reactions <sup>c</sup>	1,339	
Cytoplasmic	1,187	
Periplasmic	192	
Extracellular	8	
Transport reactions	690	
Cytoplasm to periplasm	390	
Periplasm to extracellular	298	
Cytoplasm to extracellular	2	
Gene-protein reaction associations <sup>d</sup>		
Gene associated (metabolic/transport)	1,294/625	
Spontaneous/diffusion reactions <sup>e</sup>	16/9	
Total (gene associated and no association needed)	1,310/634	(94%)
No gene association (metabolic/transport)	77/56	(6%)
Exchange reactions	304	
Metabolites		
Unique metabolites <sup>f</sup>	1,039	
Cytoplasmic	951	
Periplasm	418	
Extracellular	299	

<sup>a</sup>Overall genome coverage based on 4,453 total ORFs in *E. coli*; #F1260 contains 48% of the ORFs in *E. coli* that have been characterized experimentally (2403 ORFs)<sup>39</sup>. <sup>b</sup>Tabulated on a reaction basis, not counting outer membrane nonspecific porin transport. <sup>c</sup>Reactions can occur in or between multiple compartments and metabolites can be present in more than one compartment. <sup>d</sup>Diffusion reactions do not include facilitated diffusion reactions and are not included in this total if they can also be catalyzed by a gene product at a higher rate.

The second analysis applied the systematic computational search algorithm previously described<sup>24</sup> to the updated *E. coli* GEM MBEL979 (ref. 7) (which is similar to the *i*JR904 GEM<sup>17</sup>) to improve L-valine production. The *in silico* analysis of beneficial knockouts to divert flux toward the desired product once again resulted in a significant increase in the production of the desired metabolite over that of an existing overproducing strain, in this case, a more than a twofold increase<sup>29</sup>. Furthermore, the authors followed several additional metabolic engineering approaches to increase overproduction (that is, relieving feedback inhibition and regulation through attenuation, removing competing pathways, upregulation of primary biosynthetic pathways and overexpression of exporting machinery). When compared with each of the other individual strain modifications, the *in silico* GEM-aided interventions resulted in the greatest increase in L-valine production<sup>29</sup>. Taken together, these two studies demonstrate the broad applications for which GEMs can be used not only to design strains in a *de novo* fashion, but also to make further improvements on strains through integrating and interpreting experimental data.

Several other strain designs using *E. coli* GEMs have been reported. In a combined computational and experimental study, our group<sup>25</sup> has used the bi-level optimization algorithm OptKnock<sup>20</sup> and *i*JR904 (ref. 17) to overproduce lactate in *E. coli*. The algorithm OptKnock optimizes two objective functions, biomass formation and product secretion, to assist in the creation of strains that can couple the excretion of a desirable product



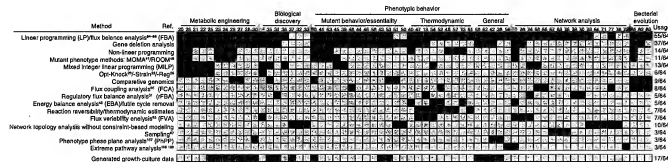
**Figure 3** Applications of the genome-scale model (GEM) of *E. coli* divided into five categories. (1) A drawing of a predicted effect from a loss-of-function mutation in a simple system is shown. Metabolic engineering studies have investigated *in silico* strain design using *E. coli* metabolic reconstructions to overproduce desired products<sup>20-30</sup>. (2) Recent studies using the reconstruction in a prospective manner have aimed to use the current biochemical and genetic information included in the metabolic network along with additional data types to drive biological discovery, such as predicting genes responsible for orphan reactions<sup>35,33,35-37</sup>. (3) Using the reconstruction in phenotypic studies, computational analyses have examined gene<sup>19,46,51,53,63</sup>, metabolite<sup>44,60</sup> and reaction<sup>39,47,48,58</sup> essentially along with considering thermodynamics<sup>19,40,47,49,52,54,55,57,61</sup> to make better predictions about the physiological state (that is, the active pathways) of the cell for a given environmental condition. (4) The *E. coli* reconstructions have been used to analyze and interpret the intrinsic properties of biological networks, one example being finding coupled reaction activities<sup>56</sup> (as shown in the drawing) across different growth conditions. (5) Using the network reconstruction, evolutionary studies have examined the cellular network in the context of adaptive evolution events<sup>81</sup>, horizontal gene transfer<sup>80,81</sup> and minimal metabolic network evolution<sup>82</sup>.

to the growth rate. Using adaptive evolution with growth-rate selection pressure, we found that the lactate-producing strains designed using OptKnock possess this growth-coupling property (as measured by growth rate, uptake and secretion rate profiles); thus, this study demonstrated the utility of adaptive evolution as a design tool<sup>87</sup>.

Additional noteworthy examples of GEM-aided design are two studies<sup>77,28</sup> demonstrating the use of GEM modeling based on iJR904 (ref. 17) in screening for genes of putative importance in succinate production. Combinatorial knockouts that were predicted to be overproducers *in silico* were experimentally verified to display the same overproducing phenotype *in vivo*. Furthermore, this method was shown to be superior to the

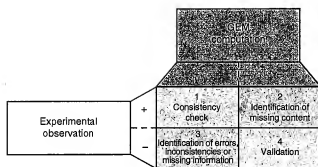
use of comparative genomics for strain design, which was also performed in one of the studies<sup>77</sup>.

A growing number of metabolic engineering studies thus demonstrate the use of GEMs to generate strain designs that are often nonintuitive and nonobvious. An excellent example of a nonintuitive strain improvement outlined in this section was when modeling was used not only to study the effect of a gene removal, but also to tune the expression of a gene to an optimally predicted level, which when surpassed, was detrimental to product formation. In this manner, genome-scale reconstructions allow the examination and simulation of metabolism as an integrated network, circumventing the possible shortcomings of methods that rely on manual



**Figure 4** Summary of the *in silico* methods used in the 64 published *E. coli* GEM studies reviewed here. This heat map characterizes the incorporation of different computational methods into studies using genome-scale models of *E. coli*<sup>20,22,26,30,40,41,45,65-67,84-86,106-108</sup>. A dark box indicates that a particular method (one method per row) was used in a corresponding study (one citation per column); the frequency of usage of a particular method is given on the right. Studies were grouped into one of five general categories, and studies examining phenotypic behavior were further divided into three subgroups. Studies that contributed new experimental growth data are also marked along the bottom offset row.

**Figure 5** Comparison of computation and experimental data: identification of agreements and disagreements. The comparison of GEM computation and organism-specific experimental measurements identifies agreements and disagreements. For comparison, phenotypic outcomes are tabulated for genetic perturbations examined in a given environment (e.g., growth or no growth) for both experiment and computation. A '+' indicates that a given phenotype is not affected by the perturbation; an '-' indicates it does. Each outcome of comparison has a different implication: 1, consistency check, a perturbation has no effect on the property being measured and modeling predicts the same; 4, validation, the perturbation affects the experimental outcome and modeling with the GEM predicts this outcome; 2, identification of missing content—when GEM modeling fails to predict the positive confirmation of the property being measured, this outcome indicates that there is missing content in the GEM and can lead to the identification of specific areas for biological discovery; 3, identification of errors, inconsistencies or missing context-specific information—a positive prediction for the measured property and an opposite experimental observation indicates a possible error in the current organism-specific knowledge or that additional context-specific information is lacking from the GEM or modeling method (e.g., transcriptional regulation).



assessment of a limited number of interactions and fail to detect non-intuitive causal interactions. With the growing availability of organism- and strain-specific GEMs, applications for designing microbial strains for industrial production are expected to continue to grow. This growth expectation is in part based on the ongoing reconstruction of additional cellular processes, such as transcriptional regulation and protein production. Computations based on genome-scale models are also beginning to influence other areas of industrial microbiology such as generation of renewable energy<sup>88–90</sup> and bioremediation<sup>89</sup>.

### GEM-driven discovery in *E. coli*

GEMs can also provide a guide to biological discovery. This capability is based on comparison of computed and actual experimental outcomes. Given the fact that BiGG knowledge bases are incomplete and that they contain gaps<sup>91</sup>, they provide a context for systematic discovery of missing information. The comparison between computation and experiments is summarized in Figure 5 highlighting how agreements and disagreements are analyzed.

The current area of most interest is direct discovery efforts toward characterizing unknown ORFs in the *E. coli* genome. Ten years after the first release of the complete genome sequence<sup>92</sup>, many unknown ORFs still exist in the *E. coli* genome (Supplementary Table 1 online), with many of these likely to encode metabolic functions. ORF discovery using GEMs also has significant potential to affect not only how new and less studied genomes are annotated, but to fill out the missing pieces of *E. coli* metabolism.

To address this challenge, researchers have developed algorithms to determine the probable gene candidates that fill knowledge gaps in the *E. coli* and other network reconstructions. These algorithms use global network topology and genomic correlations, such as genome context and protein fusion events<sup>93</sup>, as well as local network topology and/or phylogenetic profiles<sup>32,33</sup>. Similar tools have been developed that use mRNA coexpression<sup>94</sup> and which can evaluate more general metabolic pathway databases<sup>94</sup>. In addition to these network topology-based methods, an optimization-based procedure has also been developed to fill network gaps and evaluate reaction reversibility along with adding additional transport and intracellular reactions from databases of known metabolic reactions<sup>95</sup>. These studies produce specific targets for drill-down (that is, detailed biochemical characterization) experiments needed for confirmation of these computationally generated hypotheses.

Two recent studies have integrated a combined computational and experimental approach to aid the ORF discovery process in *E. coli* through

using a GEM and high-throughput phenotype data<sup>35,37</sup>. The first study (from our group<sup>35</sup>) used an iterative process in which (i) differences in modeling predictions and high-throughput growth phenotype data were identified, (ii) potential missing reactions that remedy these disagreements were algorithmically determined, (iii) bioinformatics was used to identify likely encoding ORFs and (iv) resulting targeted ORFs were cloned and experimentally characterized. Application of this process led to the functional characterization of eight ORFs involved in transport, regulatory and metabolic functions in *E. coli*<sup>35</sup>. The discovery process was aided by a high-throughput growth phenotyping analysis and the genome-wide single-gene mutant collection<sup>96</sup>, along with other characterization analyses such as targeted expression profiling.

The second GEM-based analysis resulting in ORF discovery used network topology to examine orphan reactions in the *E. coli* network (that is, reactions known to exist in *E. coli* that have not been linked to an encoding gene) identified by the previously mentioned network topology-based gap-filling algorithms<sup>32,33,93</sup>. The basic premise behind these algorithms is the utilization of an orphan reaction's network neighbors as constraints to assign metabolic function. With the resulting tentative ORF assignment, biochemical characterization on studies using genetic mutants<sup>97</sup>, analysis of growth under different substrate conditions and expression data were all used to characterize and assign function to an orphan ORF that is responsible for a metabolic conversion that has been known for 25 years<sup>97</sup>.

Further studies in this category of biological discovery applications (not focused on ORF identification) have used GEMs of *E. coli* to identify potential bottleneck reactions in the metabolic network<sup>94</sup> and as of yet uncharacterized transcription factor target interactions in *E. coli*<sup>91</sup>. The former study, targeting the elucidation of regulatory and metabolic interactions in *E. coli*, developed an iterative procedure focused on reconciling computational and experimental discrepancies stemming from high-throughput growth phenotype and gene expression data where selected expression changes were validated using RT-PCR<sup>91</sup>. With the advancement of high-throughput technologies to test the hypotheses generated from computational studies, these and similar algorithmic approaches are likely to continue to aid in the quest to achieve full functional annotation of the *E. coli* genome and its context-specific uses.

### GEM-aided phenotypic assessment

The area where the *E. coli* GEMs has been most extensively used is for the examination and quantitative interpretation of metabolic physiology for wild-type, genetically perturbed and adaptively evolved strains of *E. coli*<sup>19,38–63</sup>. These efforts have implications in both the quantitative and



qualitative understanding of physiological states of the cell. Furthermore, these efforts have examined *E. coli* physiology for a vast number of given genetic and environmental conditions and incorporation of the developed methods will have an impact on future design of biological systems and modeling approaches. A large subset of these studies of phenotypic behavior aim to use thermodynamic laws and information to refine phenotype predictions of GEMs and to incorporate metabolomic and fluxomic data into modeling<sup>19,40,47,49,52,54,55,57,61</sup>.

A set of distinct computational methods using GEMs has been developed to determine the physiological state of *E. coli* after genetic perturbations<sup>41,45,50</sup>. These studies have used <sup>13</sup>C flux measurements and growth-rate phenotype data to evaluate the predictability of the developed algorithms when compared to experimental observations. Whereas comparisons to flux data from wild-type and *E. coli* mutants reveal that the computational algorithm MOMA<sup>41</sup> provides better predictions for transient growth rates (early post-perturbation state), the algorithm ROOM (Regulatory on/off minimization)<sup>45</sup> and basic FBA was more successful in predicting final steady-state growth rates and overall lethality<sup>45</sup>. These algorithms have been used, in addition to basic FBA, for genome-wide essentiality screens, as outlined below.

A range of computational studies have sought to understand phenotypes through determining the essential genes<sup>19,46,51,53,63</sup>, metabolites<sup>44,60</sup> and reactions<sup>59,67,68,58</sup> in the *E. coli* metabolic network. A common benchmark for examining GEM predictive ability is to determine the agreement with growth phenotype data from knockout collections of *E. coli*. Such studies (e.g., refs. 19,53) will be further enabled by a comprehensive single-gene knockout library for *E. coli* recently made available<sup>65</sup>. Implications for examining network essentiality in *E. coli* include determining network essentiality in similar organisms<sup>39,48,55,58</sup>, deciphering network makeup and enzyme dispensability (that is, measures of robustness)<sup>46,58,60</sup>, aiding in metabolic network annotation, validation and refinement<sup>44</sup>, and even rescuing knockout strains through additional gene deletions<sup>63</sup>, to name a few. The predictive capability of the *E. coli* GEM, as demonstrated by these studies, has been instrumental in adapting it for different uses. One particular study examining knockout phenotypes has demonstrated that the *E. coli* GEM was able to predict the outcomes of adaptively evolved strains to a high degree (78%) when knockout *E. coli* strains were grown in several different substrate environments by examining growth rates at the beginning and end of adaptive evolution<sup>63</sup>. This study represents a demonstration of a GEM's ability to look at adaptive behavior (or 'distal' causation<sup>66</sup>), in addition to immediate behavior (or 'proximal' causation<sup>66</sup>). Predictive capability is expected to improve through examining growth behavior across a greater number of environments (additional phenotyping screens will be necessary) and with an increase of integration of additional cellular processes<sup>31</sup>. Genetic perturbations have played a key role in the study of the genotype-phenotype relationship in biology and GEMs can be used to mechanistically interpret the results and predict the outcomes of such perturbations.

Incorporating thermodynamic information into *E. coli* GEMs has shown promise in narrowing predictions of allowable physiological states in a given environment<sup>19,40,47,49,52,54,55,57,61</sup> and in identifying reactions likely to be subject to active allosteric or genetic regulation<sup>49,54</sup>. This field is progressing rapidly and should prove to increase the predictive capabilities of genome-scale modeling through the addition of governing thermodynamic physiochemical constraints. One particular analysis incorporating compound formation and reaction energies for the content of the GEM based on JH904 (ref. 17) identified reactions that are likely to be effectively irreversible for any realistic metabolite concentration<sup>54</sup>. The hypothesis was advanced that these reactions are candidates for cellular regulation in their respective pathways because enzyme regulation will likely be the dominant mechanism for control of flux through these reactions<sup>54</sup>.

The addition of thermodynamics in GEM modeling enables the analysis of metabolomic data in the context of a reconstruction. A study using high-throughput metabolomic data and GEMs proposed likely regulatory interactions by deciphering the metabolite concentrations in the context of overall network functionality<sup>49</sup>. Not only did the metabolomic data benefit computations by constraining the system using physiological measurements, but the computational predictions were also able to validate quantitative metabolomic data sets for consistency through providing a functional context to relate metabolite concentrations. This application is one example of how metabolomic data will directly influence modeling and how metabolite concentration data are likely to greatly influence future metabolic modeling owing to its intimate connection with GEM content. Similar work incorporating other quantitative values with FBA, such as metabolite concentrations<sup>57</sup> and flux ratios at branch points in metabolism<sup>56</sup>, is also appearing.

Applying a different physiochemical constraint, molecular crowding, a framework has also been developed to incorporate spatial constraints into FBA<sup>59</sup>. The functional states predicted with this method (that is, FBA with molecular crowding: FBAwMC) and the *E. coli* GEM were validated against generated growth, substrate and production rate data along with gene expression profiles and enzyme activity measures to demonstrate predictive accuracy, including substrate preferentiality, when examining growth in complex substrate environments<sup>59,62</sup>. Overall, these studies, which incorporate reaction thermodynamics and additional cellular constraints, should further narrow the range of allowable functional network states that can be based on stoichiometry alone and thus improve the utility of GEMs.

In addition to analyses on the genomic scale, several studies modeling the metabolism of *E. coli* on a smaller scale have been performed. These analyses typically use models containing ~100 reactions or less and most often focus on incorporating nonlinear analysis to understand quantitative experimental data (e.g., isotopomer modeling). With the advancement of computational power and developed platforms, the networks that can be analyzed will grow in size<sup>67</sup>. Given that the results produced from such analyses as isotopomer modeling have been shown to be highly dependent on the content of a reduced model, the logical starting point for building such models is the *E. coli* GEM<sup>67</sup>. Several noteworthy studies have been conducted with reduced models, but are not detailed here as they are outside the scope of this review.

## GEMs and network property analysis

*E. coli* is generally viewed as having the most complete characterization of any model organism<sup>68,69</sup>. Because of the incorporation of thousands of metabolite interactions with relatively high reliability (e.g., 92% of the genes included in the latest reconstruction of *E. coli*<sup>19</sup> have experimentally determined annotated functions<sup>69</sup>; Table 1), validated genome-scale reconstructions of *E. coli* have become popular resources for the analysis of various network properties<sup>64–79</sup>. The methods designed to analyze the underlying network structure of *E. coli* metabolism, some characterizing its interplay with regulation, have been developed to determine several physiological features. These features include the most probable active pathways and metabolites used under all possible growth conditions<sup>67,69,73,75</sup>, the existence of alternative optimal solutions and their physiological significance<sup>69</sup>, conserved intracellular pools of metabolites<sup>68</sup>, coupled reaction activities<sup>66</sup> and their relationship to gene co-expression<sup>77</sup>, metabolite coupling (or laboratory<sup>71</sup>), metabolite utilization<sup>72</sup>, the organization of metabolic networks<sup>64,76</sup>, strategies for *E. coli* to incorporate metabolic redundancy<sup>78</sup> and the dominant functional states of the *E. coli* network across various environments<sup>70,74,79</sup>. These findings are driven both by biased approaches using FBA and biomass objective function optimization and by unbiased approaches

such as graph-based analyses (see Fig. 4). One noteworthy study using the GEM-outlined network examined thousands of different potential growth conditions and observed a 'high-flux backbone' in *E. coli* that both carried high levels of flux across the different environmental conditions and was composed of a relatively small set of enzymatic reactions<sup>67</sup>. This result can be of practical importance for synthetic biology efforts aimed toward manipulating flux within biological systems. Furthermore, this finding was hypothesized to be a universal feature of metabolic activity in all cells and was consistent with flux measurements from <sup>13</sup>C-labeling experiments<sup>67</sup>.

The studies in this category have a common systems biology theme; namely, the development and subsequent demonstration of methods that identify sets of reactions or metabolites with correlated or coordinated functions and systematic relationships. The systems biology that these methods enable and demonstrate has potential implications for (i) antimicrobial drug-target discovery<sup>68,69</sup>, (ii) aiding the development of additional metabolic reconstructions<sup>66,68</sup>, (iii) guiding genetic manipulations<sup>66</sup>, (iv) improving metabolic engineering applications<sup>67,68</sup> and (v) increasing the general understanding of biological network behavior<sup>65,74,77</sup> and resilience<sup>78</sup>. The role that the *E. coli* GEM has taken is a comprehensive and curated set of up-to-date metabolic knowledge; thus providing a scaffold for these large-scale computations.

### GEMs and bacterial evolution

The GEMs of *E. coli* metabolism have been used to examine the process of bacterial evolution<sup>80–82</sup>. Specifically, the network reconstructions have been used to interpret adaptive evolution events<sup>81</sup>, horizontal gene transfer<sup>80,81</sup> and evolution to minimal metabolic networks<sup>82</sup>. These studies, which use the *E. coli* reconstruction as an organism-specific genetic and metabolic content database, and the corresponding GEM, have been able to provide insight into evolutionary events through combining known physiological data (e.g., in various environmental conditions) with hypotheses and *in silico* computation. Examining the evolution of minimal metabolic networks through simulation demonstrated that it was possible to predict the gene content of close relatives of *E. coli* by examining the necessity of genes and reactions in the overall context of the system functionality for a specific lifestyle<sup>82</sup>. Similarly, by re-examining network functionality in a number of different environments and through the use of comparative genomics, it has been shown that recent evolutionary events (e.g., horizontal gene transfer) probably resulted from a response to a change in environment<sup>81</sup>. Furthermore, computational analysis led to the additional conclusion that these horizontal gene transfer events are more likely if the host organism contains an enzyme that catalyzes a coupled metabolic flux related to the transferred enzyme's function<sup>80,81</sup>. Taken together, these studies demonstrate the importance of having high-quality curated reconstructions to enable studies on an organism's response to environmental changes and for understanding the fundamental forces driving bacterial evolution.

### Conclusions

Since the first review on constraint-based methods appeared in *Nature Biotechnology* in 1994 (ref. 84), the field has grown rapidly. The myriad studies described in this article highlight the rapid development and use of genome-scale reconstruction and derived computational models to address a growing spectrum of basic research and applied problems. Experience with genome-scale reconstructions has demonstrated that they are a common denominator in the systems analysis of metabolic functions. With the recognition of its basic paradigms and a growing spectrum of practical uses enabled, this field now faces several exciting challenges in three major areas, including: first, network reconstructions and the reconstruction process; second, computational BIGG query tools

(that is, modeling); and third, application to proximal and distal causation in biology.

The scope of reconstructions is bound to grow, representing more and more BIGG knowledge in the structured format of a GEM<sup>71</sup>. Growth in scope in the near term will on one front involve the transcriptional and translational machinery of bacterial cells<sup>100–102</sup>. Such an extension will enable a range of studies including the direct inclusion of proteomic data, fine graining of growth requirements and the explicit consideration of secreted protein products. Another expansion in scope in the near term is the reconstruction of the genome-scale transcriptional regulatory network (TRN). Such reconstruction at the genome-scale is now enabled by new experimental technologies, such as chromatin immunoprecipitation (ChIP)-chip<sup>103</sup>. Experimental interrogation of the currently available TRN suggests that we know about one-fourth to one-third of its content<sup>31</sup>, indicating that there is much to be discovered. Once reconstructed, the TRN will allow computational predictions of the context-specific uses of the *E. coli* genome and the responses of two-component signaling systems. Taken together, these near-term expansions in content will encompass the activity of apparently 2,000 ORFs in the *E. coli* genome.

Mid-term expansions in scope will include the growth cycle, shock responses and additional cellular functions. Such a reconstruction should eventually be a comprehensive representation of the chemical reactions and transactions enabled by *E. coli*'s gene products. Longer-term reconstruction may begin to address the three-dimensional organization of the bacterial cell. In particular, high-resolution ChIP-chip data on DNA-binding proteins could enable not only the estimation of the topological arrangement of the genome but also the elucidation of the structure of the cell wall and other cellular structures that will allow us a full three-dimensional reconstruction of *E. coli*.

We now know how to represent BIGG data in either a stoichiometric format or in the form of causal relationships (e.g., see ref. 104 from our laboratory) and how to use them to perform several lines of computational inquiries. Computational query tools of GEMs will continue to be developed. New advances will likely include modularization methods, use of fluxomic data and eventually kinetics. As the scope and content of the reconstruction grows, the need to modularize its content becomes more pressing. Fine- or coarse-grained views of cellular processes are needed for different applications. For instance, as previously mentioned, current computational limitations force the reduction in a network for the analysis of isotopomer data, and a rational way to carry out such reduction is needed. Given the systemic nature of fluxomic data and its phenotypic relevance, there is a pressing need to increase the size of the networks that can be analyzed for experimental measurement and estimation of flux states. Finally, although detailed kinetic models of microbial functions may currently be mostly of academic interest, we will most likely be able to construct them in the mid-term based on advances with metabolomic and fluxomic data, in addition to the developments that are occurring with the incorporation of thermodynamic information. Such large-scale kinetic models are likely to differ from those resulting from traditional approaches for construction of kinetic models as they come with different challenges.

As this article shows, the scope of applications of genome-scale reconstructions and GEMs is growing. Going forward, we wish to comment on three categories of applications: growth in coverage (that is, gap-filling), engineering (that is, synthetic biology) and the development of our understanding of fundamental biology (see Step 4, Fig. 1). Growth in coverage will come through discovery of missing network components. For instance, the latest metabolic reconstruction, iAF1260, contains 14% blocked reactions<sup>19</sup>. This disconnected content means that we have knowledge gaps that have arisen due to characterization of individual gene products outside the context of a given physiological function (that

is, outside a defined pathway). Metabolic profiling is one measure that will provide us with the missing upstream or downstream routes to such dead ends in the network. Also, an expansion of scope in modeling will allow further investigation of network content, such as tRNA-charging reactions that are currently in this blocked reaction set<sup>19</sup>. Furthermore, growing metabolic data suggest that we are discovering the existence of several new metabolites. Pathways that include these metabolites need to be discovered. Methods exist to compute missing pathways between molecules<sup>10</sup> that can be applied to such data. Such pathways, in turn, will lead to experimental programs to discover novel gene functions and to validate or refute the existence of such pathways. Similarly, we expect that a number of the components of TRNs are missing, such as new small RNA molecules (see **Supplementary Table 1**). Clearly, maintaining the quality control/quality assurance of such reconstructions will help in guiding us to a comprehensive genome-scale representation of all major cellular processes in bacteria at the BiGG data level of resolution that, in turn, enables GEMs of growing coverage and resolution.

Predictive models allow design. In fact, in engineering, there is 'nothing more useful than a good theory'. As this article demonstrates, genomics and high-throughput technologies have enabled the construction of predictive computational models. The scope of such predictions is limited at the moment, but with the growing scope and coverage of genome-scale reconstructions and advancements in the development of computational tools, this scope will broaden. Not only will GEMs influence design in synthetic biology, but their influence in discovery of cellular content will provide a more complete picture of the environment (that is, the parts list in the cell) in which future synthetically engineered constructs and circuits will be placed. The impact of GEMs on synthetic biology is thus likely to be notable, ranging from the provision of the cellular context of a small-scale gene circuit design to engineering of the entire genome-scale network toward fundamentally new and useful (that is, production) phenotypes.

Finally, we can speculate about the deep scientific impact that comprehensive predictive GEMs will have on our understanding of the living process. A comprehensive view of cellular functions will allow us to study the fundamental properties of both the underlying energy and information flows in living organisms. Such a view is likely to deeply affect our understanding of both distal and proximal causation in biology.

Note: Supplementary information is available on the Nature Biotechnology website.

# ACKNOWLEDGMENTS

We would like to thank Andrew Joyce, Jennifer Reed, Daniel Segre, Nathan Price, Markus Herrgard and Christian Barrett for their invaluable insight. A.M.F. is supported by National Institutes of Health R01 GM057089 grant.

Published online at <http://www.nature.com/naturebiotechnology/>  
Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>

1. Reed, J.L., Famili, I., Thiele, I. & Palsson, B.O. Towards multidimensional genome annotation. *Nat. Rev. Genet.* **7**, 130–141 (2006).
2. Palsson, B.O. *Systems Biology: Properties of Reconstructed Networks*. (Cambridge University Press, New York, 2006).
3. Price, N.D., Reed, J.L. & Palsson, B.O. Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat. Rev. Microbiol.* **2**, 895–907 (2004).
4. Frazier, M.E., Johnson, G.M., Thomassen, D.G., Oliver, C.E. & Patino, A. Realizing the potential of the Genome Revolution: The Genomes to Life Program. *Science* **300**, 290–293 (2003).
5. Reed, J.L. & Palsson, B.O. Thirteen years of building constraint-based in silico models of *Escherichia coli*. *J. Bacteriol.* **185**, 2692–2699 (2003).
6. Becker, S.A. et al. Quantitative analysis of cellular metabolism with constraint-based models: The COBRA Toolbox. *Nat. Protoc.* **2**, 727–738 (2007).
7. Lee, S.Y. et al. Systems-level analysis of genome-scale in silico metabolic models using MetaFluxNet. *Biotechnol. Bioeng.* **100**, 425–431 (2005).
8. Klamt, S., Saez-Rodriguez, J. & Gilles, E.D. Structural and functional analysis of cellular networks with CellNetAnalyzer. *BMC Syst. Biol.* **1**, 2 (2007).
9. Raman, K. & Chandra, N. PathwayAnalysr: a systems biology tool for flux analysis of metabolic pathways. Available from *Nature Precedings* <http://dx.doi.org/10.1038/npre.2008.1968.1v2> (2008).

10. Luo, R.Y., Liao, S., Zeng, S.Q., Li, Y.X. & Luo, Q.M. FluxExplorer: A general platform for modeling and analyses of metabolic networks based on stoichiometry. *Chin. Sci. Bull.* **51**, 689–696 (2006).
11. Hucka, M. et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* **19**, 524–531 (2003).
12. Majewski, R.A. & Domach, M.M. Simple constrained optimization view of acetate overflow in *E. coli*. *Biotechnol. Bioeng.* **35**, 732–738 (1990).
13. Varma, A., Boesch, B.W. & Palsson, B.O. Stoichiometric interpretation of *Escherichia coli* glycolysis under various oxygenation rates. *Appl. Environ. Microbiol.* **59**, 2465–2473 (1993).
14. Varma, A., Boesch, B.W. & Palsson, B.O. Biochemical production capabilities of *Escherichia coli*. *Biotechnol. Bioeng.* **42**, 59–73 (1993).
15. Pramanik, J. & Keasling, J.D. Stoichiometric model of *Escherichia coli* metabolism: Incorporation of growth-rate dependent biomass composition and mechanistic energy requirements. *Biotechnol. Bioeng.* **56**, 398–421 (1997).
16. Pramanik, J. & Keasling, J.D. Effect of *Escherichia coli* biomass composition on central metabolic fluxes predicted by a stoichiometric model. *Biotechnol. Bioeng.* **60**, 230–238 (1998).
17. Reed, J.L., Vo, T.D., Schilling, C.H. & Palsson, B.O. An expanded genome-scale model of *Escherichia coli* K-12 (JR994 GSMGP). *Genome Biology* **4**, R54–R54.12 (2003).
18. Edwards, J.S. & Palsson, B.O. The *Escherichia coli* MG1655 in silico metabolic genotype: Its definition, characteristics, and capabilities. *Proc. Natl. Acad. Sci. USA* **97**, 5528–5533 (2000).
19. Feist, A.M. et al. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol.* **3**, 121 (2007).
20. Burgard, A.P., Pharkya, P. & Maranas, C.D. OptKnock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.* **84**, 647–657 (2003).
21. Pharkya, P., Burgard, A.P. & Maranas, C.D. Exploring the overproduction of amino acids using the bilevel optimization framework OptKnock. *Biotechnol. Bioeng.* **84**, 887–899 (2003).
22. Pharkya, P., Burgard, A.P. & Maranas, C.D. OptStrain: a computational framework for redesign of microbial production systems. *Genome Res.* **14**, 2367–2376 (2004).
23. Alper, H., J.Y. Moyle, J.F. & Stephanopoulos, G. Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metab. Eng.* **7**, 155–164 (2005).
24. Alper, H., Miyayaka, K. & Stephanopoulos, G. Construction of lycopene overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. *Nat. Biotechnol.* **23**, 612–616 (2005).
25. Fong, S.S. et al. In silico design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnol. Bioeng.* **91**, 643–648 (2005).
26. Pharkya, P. & Maranas, C.D. An optimization framework for identifying reaction activation/inhibition or elimination candidates for overproduction in microbial systems. *Metab. Eng.* **8**, 1–13 (2006).
27. Lee, S.Y. et al. Metabolic engineering of *Escherichia coli* for enhanced production of succinic acid, based on genome comparison and in silico gene knockout simulation. *Appl. Environ. Microbiol.* **71**, 7880–7887 (2005).
28. Wang, Q., Chen, X., Yang, Y. & Zhao, X. Genome-scale in silico aided metabolic analysis and flux comparisons of *Escherichia coli* to improve succinate production. *Appl. Microbiol. Biotechnol.* **73**, 887–894 (2006).
29. Park, J.H., Lee, K.H., Kim, T.Y. & Lee, S.Y. Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proc. Natl. Acad. Sci. USA* **104**, 7797–7802 (2007).
30. Lee, K.H., Park, J.H., Kim, T.Y., Kim, H.U. & Lee, S.Y. Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Mol. Syst. Biol.* **3**, 149 (2007).
31. Covert, M.W., Knight, E.M., Reed, J.L., Herrgard, M.J. & Palsson, B.O. Integrating high-throughput and computational data elucidates bacterial networks. *Nature* **429**, 92–96 (2004).
32. Chen, L. & Vitkup, D. Predicting genes for orphan metabolic activities using phylogenetic profiles. *Genome Biol.* **7**, R17 (2006).
33. Kharchenko, P., Chen, L., Freund, Y., Vitkup, D. & Church, G.M. Identifying metabolic enzymes with multiple types of association evidence. *BMC Bioinformatics* **7**, 117 (2006).
34. Herrgard, M.J., Fong, S.S. & Palsson, B.O. Identification of genome-scale metabolic network models using experimentally measured flux profiles. *PLoS Comput. Biol.* **2**, e72 (2006).
35. Reed, J.L. et al. Systems approach to refining genome annotation. *Proc. Natl. Acad. Sci. USA* **103**, 17480–17484 (2006).
36. Satish Kumar, V., Dasika, M.S. & Maranas, C.D. Optimization based automated curation of metabolic reconstructions. *BMC Bioinformatics* **8**, 212 (2007).
37. Fuhrer, T., Chen, L., Sauer, U. & Vitkup, D. Computational prediction and experimental verification of the gene encoding the NAD<sup>+</sup>-dependent succinate semialdehyde dehydrogenase in *Escherichia coli*. *J. Bacteriol.* **189**, 8073–8079 (2007).
38. Edwards, J.S., Ibarra, R.U. & Palsson, B.O. In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat. Biotechnol.* **19**, 125–130 (2001).
39. Burgard, A.P., Vaidyanarayanan, S. & Maranas, C.D. Minimal reaction sets for *Escherichia coli* metabolism under different growth requirements and uptake environments. *Biotechnol. Prog.* **17**, 791–797 (2001).

40. Beard, D.A., Liang, S.D. & Qian, H. Energy balance for analysis of complex metabolic networks. *Biophys. J.* **83**, 79–86 (2002).
41. Segre, D., Vitkup, D. & Church, G.M. Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. USA* **99**, 15112–15117 (2002).
42. Ibarra, R.U., Edwards, J.S. & Palsson, B.O. *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* **410**, 186–189 (2002).
43. Fong, S.S. & Palsson, B.O. Metabolic gene-deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes. *Nat. Genet.* **36**, 1056–1058 (2004).
44. Imielinski, W., Sella, G., Hatzias, A. & Rubin, H. Investigating metabolite essentiality through genome-scale analysis of *Escherichia coli* production capabilities. *Bioinformatics* **21**, 2008–2016 (2005).
45. Shlomi, T., Berkman, O. & Ruppin, E. Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proc. Natl. Acad. Sci. USA* **102**, 7695–7700 (2005).
46. Ghim, C.M., Goh, K.J. & Kahng, B. Lethality and synthetic lethality in the genome-wide metabolic network of *Escherichia coli*. *J. Theor. Biol.* **237**, 401–411 (2005).
47. Henry, C.S., Jankowski, M.D., Broadbelt, L.J. & Hatzimanikatis, V. Genome-scale thermodynamic analysis of *Escherichia coli* metabolism. *Biomol. J.* **90**, 1453–1461 (2006).
48. Samal, A. et al. Low degree metabolites explain essential reactions and enhance modularity in biological networks. *BMC Bioinformatics* **7**, 118 (2006).
49. Kümmel, A., Panke, S. & Heinemann, M. Putative regulatory sites unraveled by network-embedded thermodynamic analysis of metabolite data. *Mol. Syst. Biol.* **2**, 2006.0034 (2006).
50. Wunderlich, Z. & Mirny, L.A. Using the topology of metabolic networks to predict viability of mutant strains. *Biophys. J.* **91**, 2304–2311 (2006).
51. Gardes, S. et al. Essential genes in metabolic maps. *Curr. Opin. Biotechnol.* **17**, 448–456 (2006).
52. Kümmel, A., Panke, S. & Heinemann, M. Systematic assignment of thermodynamic constraints in metabolic network models. *BMC Bioinformatics* **7**, 512 (2006).
53. Joyce, A.R. et al. Experimental and computational assessment of conditionally essential genes in *Escherichia coli*. *J. Bacteriol.* **188**, 8259–8271 (2006).
54. Henry, C.S., Broadbelt, L.J. & Hatzimanikatis, V. Thermodynamics-based reaction flux analysis. *Biophys. J.* **92**, 1792–1805 (2007).
55. Ederer, M. & Gilles, E.D. Thermodynamically feasible kinetic models of reaction networks. *Biophys. J.* **92**, 1846–1857 (2007).
56. Choi, H.S., Kim, T.Y., Lee, D.Y. & Lee, S.Y. Incorporating metabolic flux ratios into constraint-based flux analysis by using artificial metabolites and converging ratio determinants. *J. Biotechnol.* **129**, 696–705 (2007).
57. Hoppe, A., Hoffmann, S. & Holzhüter, H.G. Including metabolite concentrations into flux balance analysis: thermodynamic realizability as a constraint on flux distributions in metabolic networks. *BMC Syst. Biol.* **1**, 23 (2007).
58. Guimera, R., Sales-Pardo, M. & Amaral, L.A. A network-based method for target selection in metabolic networks. *Bioinformatics* **23**, 1616–1622 (2007).
59. Beg, Q.K. et al. Intracellular crowding defines the mode and sequence of substrate uptake by *Escherichia coli* and constrains its metabolic activity. *Proc. Natl. Acad. Sci. USA* **104**, 12663–12668 (2007).
60. Kim, P.J. et al. Metabolic essentiality elucidates robustness of *Escherichia coli* metabolism. *Proc. Natl. Acad. Sci. USA* **104**, 13638–13642 (2007).
61. Warren, P.B. & Jones, J.L. Duality, thermodynamics, and the linear programming problem in constraint-based models of metabolism. *Phys. Rev. Lett.* **99**, 108101 (2007).
62. Vazquez, A. et al. Impact of the solvent capacity constraint on *E. coli* metabolism. *BMC Syst. Biol.* **2**, 7 (2008).
63. Motter, A.E., Gulloche, N., Almeida, E. & Barabási, A.L. Predicting synthetic rescues in metabolic networks. *Mol. Syst. Biol.* **4**, 168 (2008).
64. Gagneur, J., Jackson, D.B. & Casari, G. Hierarchical analysis of dependency in metabolic networks. *Bioinformatics* **19**, 1027–1034 (2003).
65. Mahadevan, R. & Schilling, C.H. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab. Eng.* **5**, 264–276 (2003).
66. Burgard, A.P., Nikolaev, E.V., Schilling, C.H. & Maranas, C.D. Flux coupling analysis of genome-scale metabolic network reconstructions. *Genome Res.* **14**, 301–312 (2004).
67. Almeida, E., Kovacs, B., Visek, T., Olvai, Z.N. & Barabási, A.L. Global organization of metabolic fluxes in the bacterium *Escherichia coli*. *Nature* **427**, 839–843 (2004).
68. Nikolaev, E.V., Burgard, A.P. & Maranas, C.D. Elucidation and structural analysis of conserved pools for genome-scale metabolic reconstructions. *Biophys. J.* **88**, 37–49 (2005).
69. Almeida, E., Olvai, Z.N. & Barabási, A.L. The activity reaction core and plasticity of metabolic networks. *PLoS Comput. Biol.* **1**, e68 (2005).
70. Barrett, C.J., Herring, C.P., Reed, J.L. & Palsson, B.O. The global transcriptional regulatory network for metabolism in *Escherichia coli* attains few dominant functional states. *Proc. Natl. Acad. Sci. USA* **102**, 19103–19108 (2005).
71. Becker, S.A., Price, N.D. & Palsson, B.O. Metabolite coupling in genome-scale metabolic networks. *BMC Bioinformatics* **7**, 111 (2006).
72. Imielinski, W., Sella, G., Rubin, H. & Hatzias, A. Systematic analysis of conservation relations in *Escherichia coli* genome-scale metabolic network reveals novel growth media. *Biophys. J.* **90**, 2659–2672 (2006).
73. Seasley, J.E. & Pflans, F. Recovering metabolic pathways via optimization. *Bioinformatics* **23**, 92–98 (2007).
74. Shlomi, T., Eisenberg, Y., Sharan, R. & Ruppin, E. A genome-scale computational study of the interplay between transcriptional regulation and metabolism. *Mol. Syst. Biol.* **3**, 101 (2007).
75. Almeida, E. Optimal flux patterns in cellular metabolic networks. *Chaos (Woodbury, N.Y.)* **17**, 025107 (2007).
76. Sales-Pardo, M., Guimera, R., Moreira, A.A. & Amaral, L.A. Extracting the hierarchical organization of complex systems. *Proc. Natl. Acad. Sci. USA* **104**, 15229–15239 (2007).
77. Notebaart, R.A., Teusink, B., Sieszen, R.J. & Papp, B. Co-regulation of metabolic genes is better explained by flux coupling than by network distance. *PLoS Comput. Biol.* **4**, e26 (2008).
78. Mahadevan, R. & Lovley, D.R. The degree of redundancy in metabolic genes is linked to mode of metabolism. *Biophys. J.* **94**, 1216–1220 (2008).
79. Samal, A. & Jain, S. The regulatory network of *E. coli* metabolism as a Boolean dynamical system exhibits both homeostasis and flexibility of response. *BMC Syst. Biol.* **2**, 21 (2008).
80. Pal, C., Papp, B. & Lercher, M.J. Horizontal gene transfer depends on gene content of the host. *Bioinformatics* **21** Suppl 2, i222–i223 (2005).
81. Pal, C., Papp, B. & Lercher, M.J. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat. Genet.* **37**, 1372–1375 (2005).
82. Pal, C. et al. Chance and necessity in the evolution of minimal metabolic networks. *Nature* **440**, 667–670 (2006).
83. Schuetz, R., Kuepfer, L. & Sauer, U. Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol. Syst. Biol.* **3**, 119 (2007).
84. Varma, A. & Palsson, B.O. Metabolic flux balancing: basic concepts, scientific and practical use. *Nat. Biotechnol.* **12**, 994–998 (1994).
85. Edwards, J.S., Ramakrishna, R., Schilling, C.H. & Palsson, B.O. Metabolic flux balance analysis. In *Metabolic Engineering* (eds. Lee, S.Y. & Papoutsakis, E.T.) (Marcel Dekker, New York, 1999).
86. Kauffman, K.J., Prakash, P. & Edwards, J.S. Advances in flux balance analysis. *Curr. Opin. Biotechnol.* **14**, 491–496 (2003).
87. Fraser-Liggett, C.M. Insights on biology and evolution from microbial genome sequencing. *Genome Res.* **15**, 1603–1610 (2005).
88. Ro, C., Regenstein, B., Forster, J. & Nielsen, J. In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metab. Eng.* **8**, 102–111 (2006).
89. Mahadevan, R. et al. Characterization of metabolism in the Fe(II)-reducing organism *Geobacter sulfurreducens* by constraint-based modeling. *Appl. Environ. Microbiol.* **72**, 1558–1568 (2006).
90. Feist, A.M., Scholten, J.C.M., Palsson, B.O., Brockman, F.J. & Ideker, T. Modeling methanogenesis with a genome-scale metabolic reconstruction of *Methanosarcina barkeri*. *Mol. Syst. Biol.* **2**, 1–14 (2006).
91. Breitling, R., Vitkup, D. & Barrett, M.P. New surveyor tools for charting microbial metabolisms. *Nat. Rev. Microbiol.* **6**, 156–161 (2008).
92. Blattner, F.R. et al. The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1474 (1997).
93. Kharochov, P., Vitkup, D. & Church, G.M. Filling gaps in a metabolic network using expression information. *Bioinformatics* **20** Suppl 1, i178–i185 (2004).
94. Green, M.L. & Karp, P.D. A Bayesian method for identifying missing enzymes in predicted metabolic pathway databases. *BMC Bioinformatics* **5**, 76 (2004).
95. Baba, T. et al. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
96. Mayr, E. *This is Biology: The Science of the Living World*. (Belknap Press of Harvard University Press, Cambridge, MA, 1997).
97. Suthers, P.F. et al. Metabolic flux elucidation for large-scale models using 13C labeled isotopes. *Metab. Eng.* **9**, 387–405 (2007).
98. Jansson, P., Goldovsky, L., Kunin, V., Darsent, N. & Ouzounis, C.A. Genome coverage, literally speaking: The challenge of annotating 200 genomes with 4 million publications. *EMBO Rep.* **6**, 397–399 (2005).
99. Riley, M. et al. *Escherichia coli* K-12: a cooperatively developed annotation snapshot-2005. *Nucleic Acids Res.* **34**, 1–9 (2006).
100. Allen, T.E. & Palsson, B.O. Sequence-based analysis of metabolic demands for protein synthesis in prokaryotes. *J. Theor. Biol.* **220**, 1–18 (2003).
101. Moha, A. & Hatzimanikatis, V. An algorithmic framework for genome-wide modeling and analysis of translation networks. *Biophys. J.* **90**, 1136–1146 (2006).
102. Thomas, R., Paredes, C.J., Mehrotra, S., Hatzimanikatis, V. & Papoutsakis, E.T. A model-based optimization framework for the inference of regulatory interactions using time-course DNA microarray expression data. *BMC Bioinformatics* **8**, 228 (2007).
103. Lee, T.I. et al. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**, 799–804 (2002).
104. Gianchandani, E.P., Papin, J.A., Price, N.D., Joyce, A.R. & Palsson, B.O. Matrix formalization of states of transcriptional regulatory systems. *PLoS Comput. Biol.* **2**, e101 (2006).
105. Li, C. et al. Computational discovery of biochemical routes to specialty chemicals. *Chem. Eng. Sci.* **59**, 5051–5060 (2004).
106. Edwards, J.S., Ramakrishna, R. & Palsson, B.O. Characterizing the metabolic phenotype of a phase plane analysis. *Biotechnol. Bioeng.* **77**, 27–36 (2002).
107. Schuster, S. & Hilgert, C. On elementary flux modes in biochemical reaction systems at steady state. *J. Biol. Syst.* **2**, 165–182 (1994).
108. Schilling, C.H., Letscher, D. & Palsson, B.O. Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *J. Theor. Biol.* **203**, 229–248 (2000).

# The *Escherichia coli* MG1655 *in silico* metabolic genotype: Its definition, characteristics, and capabilities

J. S. Edwards\* and B. O. Palsson†

Department of Bioengineering, University of California, San Diego, La Jolla, CA 92093-0412

Communicated by Yuan-Cheng B. Fung, University of California, San Diego, La Jolla, CA, March 3, 2000 (received for review October 14, 1999)

The *Escherichia coli* MG1655 genome has been completely sequenced. The annotated sequence, biochemical information, and other information were used to reconstruct the *E. coli* metabolic map. The stoichiometric coefficients for each metabolic enzyme in the *E. coli* metabolic map were assembled to construct a genome-specific stoichiometric matrix. The *E. coli* stoichiometric matrix was used to define the system's characteristics and the capabilities of *E. coli* metabolism. The effects of gene deletions in the central metabolic pathways on the ability of the *in silico* metabolic network to support growth were assessed, and the *in silico* predictions were compared with experimental observations. It was shown that based on stoichiometric and capacity constraints the *in silico* analysis was able to qualitatively predict the growth potential of mutant strains in 86% of the cases examined. Herein, it is demonstrated that the synthesis of *in silico* metabolic genotypes based on genomic, biochemical, and strain-specific information is possible, and that systems analysis methods are available to analyze and interpret the metabolic phenotype.

bioinformatics | metabolism | genotype-phenotype relation | flux balance analysis

The complete genome sequence for a number of microorganisms has been established (The Institute for Genomic Research at [www.tigr.org](http://www.tigr.org)). The genome sequencing efforts and the subsequent bioinformatic analyses have defined the molecular "parts catalogue" for a number of living organisms. However, it is evident that cellular functions are multigenic in nature, thus one must go beyond a molecular parts catalogue to elucidate integrated cellular functions based on the molecular cellular components (1). Therefore, to analyze the properties and the behavior of complex cellular networks, one needs to use methods that focus on the systemic properties of the network. Approaches to analyze, interpret, and ultimately predict cellular behavior based on genomic and biochemical data likely will involve bioinformatics and computational biology and form the basis for subsequent bioengineering analysis.

In moving toward the goal of developing an integrated description of cellular processes, it should be recognized that there exists a history of studying the systemic properties of metabolic networks (2) and many mathematical methods have been developed to carry out such studies. These methods include approaches such as metabolic control analysis (3, 4), flux balance analysis (FBA) (5–7), metabolic pathway analysis (8–11, 69), cybernetic modeling (12), biochemical systems theory (13), temporal decomposition (14), and so on. Although many mathematical methods and approaches have been developed, there are few comprehensive metabolic systems for which detailed kinetic information is available and where such detailed analysis can be carried out (see refs. 15–17 for a few noteworthy exceptions).

To analyze, interpret, and predict cellular behavior, each individual step in a biochemical network must be described, normally with a rate equation that requires a number of kinetic

constants. Unfortunately, it currently is not possible to formulate this level of description of cellular processes on a genome scale. The kinetic parameters cannot be estimated from the genome sequence and these parameters are not available in the literature. In the absence of kinetic information, it is, however, still possible to assess the theoretical capabilities of one integrated cellular process, namely metabolism, and examine the feasible metabolic flux distributions under a steady-state assumption. The steady-state analysis is based on the constraints imposed on the metabolic network by the stoichiometry of the metabolic reactions, which basically represent mass balance constraints. The steady-state analysis of metabolic networks based on the mass balance constraints is known as FBA (7, 18, 19). This analysis differs from detailed kinetic modeling of cellular processes, in that it does not attempt to predict the exact behavior of metabolic networks. Rather it uses known constraints on the integrated function of multiple enzymes to separate the states that a system can reach from those that it cannot. Then within the domain of allowable behavior one can study the genotype-phenotype relation, such as the stoichiometric optimal growth performance in a defined environment.

In this manuscript, we have used the biochemical literature, the annotated genome sequence data, and strain-specific information, to formulate an organism scale *in silico* representation of the *Escherichia coli* MG1655 metabolic capabilities. FBA then was used to assess metabolic capabilities subject to these constraints leading to qualitative predictions of growth performance.

## Materials and Methods

**Definition of the *E. coli* MG1655 Metabolic Map.** An *in silico* representation of *E. coli* metabolism has been constructed. We have used the biochemical literature (20), genomic information (21), and the metabolic databases (22–24). Because of the long history of *E. coli* research, there was biochemical or genetic evidence for every metabolic reaction included in the *in silico* representation, and in most cases, there was both genetic and biochemical evidence (Table 1). The complete list of genes included in the *in silico* analysis is shown in Table 1, and the metabolic reactions catalyzed by these genes can be found on the web (<http://gcr.ucsd.edu/downloads.html>). The stoichiometric coefficients for each metabolic reaction within this list were used to form the stoichiometric matrix *S*.

**Determining the Capabilities of the *E. coli* Metabolic Network.** The theoretical metabolic capabilities of *E. coli* were assessed by FBA

Abbreviations: FBA, flux balance analysis; LP, linear programming; TCA, tricarboxylic acid; PPP, pentose phosphate pathway.

\*Present address: Department of Genetics, Harvard Medical School, Boston, MA 02115.

†To whom reprint requests should be addressed. E-mail: [palsson@ucsd.edu](mailto:palsson@ucsd.edu).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.



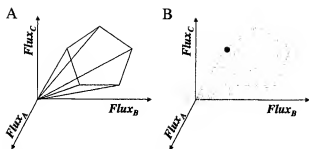


Fig. 1. The feasible solution set for a hypothetical metabolic reaction network. (A) The steady-state operation of the metabolic network is restricted to the region within a cone, defined as the feasible set (8). The feasible set contains all flux vectors that satisfy the physicochemical constraints (Eqs. 1 and 2). Thus, the feasible set defines the capabilities of the metabolic network. All feasible metabolic flux distributions lie within the feasible set, and (B) in the limiting case, where all constraints on the metabolic network are known, such as the enzyme kinetics and gene regulation, the feasible set may be reduced to a single point. This single point must lie within the feasible set.

imal metabolic fluxes in the transport reactions. The intersection of the nullspace and the region defined by the linear inequalities formally defined a region in flux space that we will refer to as the feasible set. The feasible set defined the capabilities of the metabolic network subject to the subset of cellular constraints, and all feasible metabolic flux distributions lie within the feasible set (see Fig. 1). However, every vector  $v$  within the feasible set is not reachable by the cell under a given condition because of other constraints not considered in the analysis (i.e., maximal internal fluxes and gene regulation). The feasible set can be further reduced by imposing additional constraints, and if all of the necessary details to describe metabolic dynamics are known, then the feasible set may reduce to a small region or even a single point (see Fig. 1).

For the analysis presented herein, we defined  $\alpha_i = 0$  for irreversible internal fluxes, and  $\alpha_i = -\infty$  for reversible internal fluxes. The reversibility of the metabolic reactions was determined from the biochemical literature and is identified for each reaction on the web site. The transport flux for inorganic phosphate, ammonia, carbon dioxide, sulfate, potassium, and sodium was unconstrained ( $\alpha_i = -\infty$  and  $\beta_i = \infty$ ). The transport flux for the other metabolites, when available in the *in silico* medium, was constrained between zero and the maximal level

( $0 < v_i < v_i^{\max}$ ). However, when the metabolite was not available in the medium, the transport flux was constrained to zero. The transport flux for metabolites that were capable of leaving the metabolic network (i.e., acetate, ethanol, lactate, succinate, formate, pyruvate, etc.) always was unconstrained in the outward direction.

A particular metabolic flux distribution within the feasible set was found by using linear programming (LP). A commercially available LP package was used (LINDO, Lindo Systems, Chicago). LP identified a solution that minimized a particular metabolic objective (subject to the imposed constraints) (5, 25, 26), and was formulated as shown. Minimize  $-Z$ , where

$$Z = \sum c_i v_i = (c \cdot v). \quad [3]$$

The vector  $c$  was used to select a linear combination of metabolic fluxes to include in the objective function (27). Herein,  $c$  was defined as the unit vector in the direction of the growth flux, and the growth flux was defined in terms of the biosynthetic requirements:

$$\sum_{all\ m} d_m \cdot X_m \xrightarrow{v_{growth}} \text{Biomass}, \quad [4]$$

where  $d_m$  is the biomass composition of metabolite  $X_m$  (defined from the literature; ref. 28), and the growth flux is modeled as a single reaction that converts all of the biosynthetic precursors into biomass.

## Results

FBA was used to examine the change in the metabolic capabilities caused by gene deletions. To simulate a gene deletion, the flux through the corresponding enzymatic reaction was restricted to zero. Genes that code for isozymes or genes that code for components of same enzyme complex were simultaneously removed (i.e., *aceEF*, *sucCD*). The optimal value of the objective ( $Z_{mutant}$ ) was compared with the "wild-type" objective ( $Z$ ) to determine the systemic effect of the gene deletion. The ratio of optimal growth yields ( $Z_{mutant}/Z$ ) was calculated (Fig. 2).

**Gene Deletions.** *E. coli* MG1655 *in silico* was subjected to deletion of each individual gene product in the central metabolic pathways [glycolysis, pentose phosphate pathway (PPP), tricarboxylic acid (TCA) cycle, respiration processes], and the maximal ca-

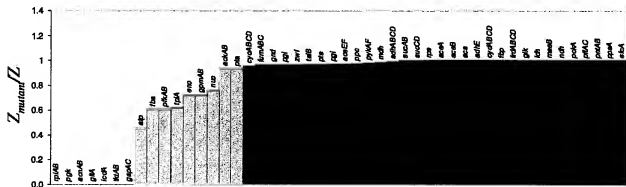


Fig. 2. Gene deletions in *E. coli* MG1655 central intermediary metabolism; maximal biomass yields on glucose for all possible single gene deletions in the central metabolic pathways. The optimal value of the mutant objective function ( $Z_{mutant}$ ) compared with the "wild-type" objective function ( $Z$ ), where  $Z$  is defined in Eq. 3. The ratio of optimal growth yields ( $Z_{mutant}/Z$ ). The results were generated in a simulated aerobic environment with glucose as the carbon source. The transport fluxes were constrained as follows:  $R_{glucose} = 10 \text{ mmol/g-dry weight (DW) per h}$ ;  $R_{oxygen} = 15 \text{ mmol/g-DW per h}$ . The maximal yields were calculated by using FBA with the objective of maximizing growth. The biomass yields are normalized with respect to the results for the full metabolic genotype. The yellow bars represent gene deletions that reduced the maximal biomass yield to less than 95% of the *in silico* wild type.

ability of each *in silico* mutant metabolic network to support growth was assessed with FBA. The simulations were performed under an aerobic growth environment on minimal glucose medium.

The results identified the essential (required for growth) central metabolic genes (Fig. 2). For growth on glucose, the essential gene products were involved in the three-carbon stage of glycolysis, three reactions of the TCA cycle, and several points within the PPP. The remainder of the central metabolic genes could be removed and *E. coli in silico* maintained the potential to support cellular growth. This result was related to the interconnectivity of the metabolic reactions. The *in silico* gene deletion results suggest that a large number of the central metabolic genes can be removed without eliminating the capability of the metabolic network to support growth under the conditions considered.

**Are the *in Silico* Redundancy Results Consistent with Mutant Data?** The *in silico* gene deletion study results were compared with growth data from known mutants. The growth characteristics of a series of *E. coli* mutants on several different carbon sources were examined and compared with the *in silico* deletion results (Table 2). From this analysis, 86% (68 of 79 cases) of the *in silico* predictions were consistent with the experimental observations.

**How Are Cellular Fluxes Redistributed?** The potential of many *in silico* deletion strains to support growth led to questions regarding how the *E. coli* metabolic genotype deals with the loss of metabolic functions. The answer involves the degree of stoichiometric connectivity of key metabolites. For illustration, the flux redistributions to optimally support growth of a single mutant and a double mutant were investigated.

The optimal metabolic flux distribution for the *in silico* wild type was calculated (Fig. 3). The constraints used in the LP problem are defined in the figure legend. The *in silico* results suggest that optimally the oxidative branch of the PPP was used to generate a large fraction of the NADPH (66% *in silico*: 20–50% reported in the literature, ref. 29), and the TCA cycle produced NADH. The optimal flux distribution also suggested that the majority of the high-energy phosphate bonds were generated via oxidative phosphorylation and acetate secretion because of limitations of the oxygen supply.

The *in silico* gene deletion results predicted that the optimal biomass yield of the *zwf* (glucose-6-phosphate dehydrogenase) *in silico* strain was slightly less than the wild type. The optimal flux distribution of the *zwf in silico* strain (Fig. 2) was calculated, and the NADPH was optimally generated through the transhydrogenase reaction and an elevated TCA cycle flux. The PPP biosynthetic precursors were generated in the nonoxidative branch. This metabolic flux rerouting resulted in an optimal biomass yield that was 99% of the *in silico* wild type.

The transhydrogenase (*pnt*) also was deleted *in silico*, creating an *in silico* double deletion mutant and eliminating an alternate source of NADPH. The double mutant still maintained growth potential. The optimal flux distribution (Fig. 2) used the isocitrate dehydrogenase and the malic enzyme to produce NADPH. The optimal biomass yield of the double mutant was 92% of the *in silico* wild type. The FBA results were consistent with the experimental observations that the *zwf* strain (30) and the *pnt* strain (29) are able to grow at near wild-type yields. Furthermore, the *zwf pnt* double mutant strain also has been shown to grow ( $\mu_{\text{mutant}}/\mu_{\text{wild type}} = 57\%$ ) (29).

## Discussion

Extensive information about the molecular composition and function of several single-cellular organisms has become available. A next important step will be to incorporate the available information to generate whole-cell models with interpretative

**Table 2. Comparison of the predicted mutant growth characteristics from the gene deletion study to published experimental results with single mutants**

Gene	glc	gl	succ	ac	Reference
<i>aceA</i>	+/+		+/+	—/—	(58)
<i>aceB</i>				—/—	(58)
<i>aceEF*</i>	+/+				(60)
<i>ackA</i>				+/+	(61)
<i>acn</i>	—/—			—/—	(58)
<i>acs</i>				+/+	(61)
<i>cyd</i>	+/+				(62)
<i>cyo</i>	+/+				(62)
<i>eno*</i>	+/+	+/+	—/—	—/—	(30)
<i>fab*</i>	+/+	—/—	—/—	—/—	(30)
<i>fbp</i>	+/+	—/—	—/—	—/—	(30)
<i>frd</i>	+/+		+/+	+/+	(60)
<i>gap</i>	—/—	—/—	—/—	—/—	(30)
<i>gik</i>	+/+				(30)
<i>glTA</i>	—/—			—/—	(58)
<i>gnd</i>	+/+				(30)
<i>idh</i>	—/—			—/—	(58)
<i>mdh*</i>	+/+	+/+	+/+		(63)
<i>ndh</i>	+/+	+/+			(59)
<i>nuo</i>	+/+	+/+			(59)
<i>pfk*</i>	+/+	+/+	+/+		(30)
<i>pgi*</i>	+/+	+/+	+/+		(30)
<i>pgk</i>	—/—	—/—	—/—	—/—	(30)
<i>pgl</i>	+/+				(30)
<i>pntAB</i>	+/+	+/+	+/+		(29)
<i>ppc*</i>	±/±	—/+	+/+		(63, 64)
<i>pta</i>				+/+	(61)
<i>pts</i>	+/+				(30)
<i>pyk</i>	+/+				(30)
<i>rpi</i>	—/—	—/—	—/—	—/—	(30)
<i>sdhABCD</i>	+/+		—/—	—/—	(58)
<i>sucAB</i>	+/+		—/+	—/+	(60)
<i>tktAB</i>	—/—				(30)
<i>tpi**</i>	+/+	—/—	—/—	—/—	(30)
<i>unc</i>	+/+		±/±	—/—	(66–68)
<i>zwf</i>	+/+	+/+	+/+		(30)

Results are scored as + or — meaning growth or no growth determined from *in vivo* *in silico* data. The ± indicates that suppressor mutations have been observed that allow the mutant strain to grow. In 68 of 79 cases the *in silico* behavior is the same as the experimentally observed behavior. glc, glucose; ac, acetate; gl, glycerol; succ, succinate.

\*The *in vivo* *aceAE* strain is able to grow under anaerobic growth conditions by using the pyruvate formate lyase.

†The *in silico* *pfk* strain is able to grow by increasing the PPP flux  $\sim 5\times$  and using the *ppg* gene product to overcome PEP deficiency.

‡The *in silico* *pgi* strain is unable to grow with glycerol or succinate as the carbon source because it is unable to synthesize glycogen and one carbohydrate component in the lipopolysaccharide. These are likely nonessential components of the biomass.

§The growth on glycerol and glucose is possible through the utilization of the glyoxylate bypass. Constitutive mutations in the glyoxylate bypass can suppress the *ppc* phenotype.

¶The *in silico* *eno* strain is able to grow by the synthesis and degradation of serine.

‡‡There is evidence that *iba* has an inhibitory effect on stable RNA synthesis (65). Such an inhibition cannot be predicted by FBA.

••The inability of *tpi* mutants to grow on glucose may be related to the accumulation of dihydroxyacetone phosphate, which leads to the formation of the bactericidal compound methylglyoxal (30).

††Very slow growth on glycerol and succinate.

and predictive capability. Herein, we have taken a step in that direction by using a set of constraints on cellular metabolism on the whole-cell level to analyze the metabolic capabilities of the



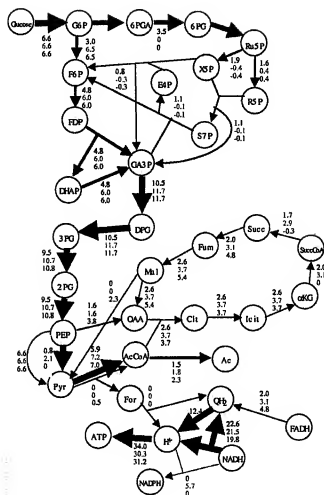


Fig. 3. Rerouting of metabolic fluxes. (Black) Flux distribution for the complete gene set. (Red) *zwf* mutant. Biomass yield is 99% of the results for the full metabolic genotype. (Blue) *zwf pnt* mutant. Biomass yield is 92% of the results for the full metabolic genotype (see text). The solid lines represent enzymes that are being used, with the corresponding flux value noted. The fluxes [substrates converted/h per g-dry weight (DW)] were calculated by using FBA with the input parameters of glucose uptake rate ( $\beta_{\text{glucose}} = 6.6$  mmol glucose/h per g-DW) and oxygen uptake rate ( $\beta_{\text{oxygen}} = 12.4$  mmol oxygen/h per g-DW) (41).

extensively studied bacterium *E. coli*. We have calculated the optimal metabolic network utilization with a FBA. The *in silico* results, based only on stoichiometric and capacity constraints, were consistent with experimental data for the wild type and many of the mutant strains examined.

The construction of comprehensive *in silico* metabolic maps provided a framework to study the consequences of alterations in the genotype and to gain insight into the genotype-phenotype relation. The stoichiometric matrix and FBA were used to analyze the consequences of the loss of a gene product function on the metabolic capabilities of *E. coli*. The results demonstrated an important property of the *E. coli* metabolic network, namely that there are relatively few critical gene products in central metabolism. The nonessential genes in several organisms have been found experimentally on a genome scale (31, 32), which opens up the opportunity to critically test the *in silico* predictions. The *in silico* analysis also suggests that although the ability to grow in one defined environment is only slightly altered the ability to adjust to different environments may be diminished

(33). Therefore, the *in silico* analysis provides a methodology for relating the specific biochemical function of the metabolic enzymes to the integrated properties of the metabolic network.

The *in silico* analysis presented herein is not the typical metabolic modeling; more appropriately, the analysis can be thought of as a constraining approach. This approach defines the "best" the cell can do and identifies what the cell cannot do, rather than attempting to predict how the cell actually will behave under a given set of conditions. To accomplish this, we have used a set of physicochemical constraints for which there is reliable information available, in particular the stoichiometric properties. FBA does not directly consider regulation or the regulatory constraints on the metabolic network.

The results of FBA can be interpreted in a qualitative or a quantitative sense. At the first level we can ask whether a cell is able to grow under given circumstances and how a loss of the function of a gene product influences this ability. The results presented herein fall into this category. Quantitative predictions would hold true if the cell optimized its growth under the growth conditions considered. Therefore, when applying LP to predict quantitatively the optimal metabolic pathway utilization, it is assumed that the cell has found an "optimal solution" for survival through natural selection, and we have equated survival with growth. Although *E. coli* may grow optimally in defined media, one should not expect that optimizing growth is the governing objective of the cell under all growth conditions. For example, the regulatory mechanisms can only evolve to stoichiometric optimality in a condition to which the cell has been exposed. Furthermore, the growth behavior of mutant strains is unlikely to be optimal. However, FBA can still be used to delineate the metabolic capabilities of mutant cells based on constraining features, because both wild-type and mutant cells must obey the physicochemical constraints imposed.

The constraints on the system accurately reflect the steady-state capabilities of the metabolic network, but does the calculated optimal flux vector in the feasible set accurately reflect the behavior of the actual metabolic network? It has been shown that in a minimal media the metabolic behavior of wild-type *E. coli* is consistent with stoichiometric optimality (34). Furthermore, more detailed and critical experimental results are consistent with the hypothesis that *E. coli* does optimize its growth in acetate or succinate minimal media (33). Taken together these results call for critical experimental investigation to evaluate the hypothesis that stoichiometric and capacity constraints are the principal constraints that limit *E. coli* maximal growth. Even though growth and metabolic behavior in minimal media are consistent with FBA results, one still must determine the generality of optimal performance. The call for critical experimentation is particularly timely, given the increasing number of genome scale measurements that are now possible through two-dimensional gels (35, 36) and DNA array technology (37, 38). Furthermore, the ability to precisely remove ORFs can be used to design critical experiments (39). The *in silico* model can be used to choose the most informative knockouts and to design growth experiments with the knockouts.

At the present time, the annotation of the *E. coli* genome is incomplete, and about one-third of its ORFs do not have a functional assignment. Thus, the metabolic genotype studied here may lack some metabolic capabilities that *E. coli* possesses. The biochemical literature also was used to define the *in silico* metabolic genotype, and given the long history of *E. coli* metabolic research (20), a large percentage of the *E. coli* metabolic capabilities likely have been identified. However, if additional metabolic capabilities are discovered (40), the *E. coli* stoichiometric matrix can be updated, leading to an iterative model building process. Additionally, the *in silico* analysis can help identify missing or incorrect functional assignments by

identifying sets of metabolic reactions that are not connected to the metabolic network by the mass balance constraints.

The ability to analyze, interpret, and ultimately predict cellular behavior has been a long sought-after goal. The genome sequencing projects are defining the molecular components within the cell, and describing the integrated function of these molecular components will be a challenging task. The results presented herein suggest that it may be possible to analyze cellular metabolism based on a subset of the constraining features. Continued prediction and experimental verification will be an

integral part in the further development of *in silico* strains. Deciphering the complex relation between the genotype and the phenotype will involve the biological sciences, computer science, and quantitative analysis, all of which must be included in the bioengineering of the 21st century.

We thank Ramprasad Ramakrishna, George Church, and Christophe Schilling for critical advice and input. National Institutes of Health Grant GM 57089 and National Science Foundation Grant MCB 9873384 supported this research.

- Weng, G., Bhalla, U. S. & Iyengar, R. (1999) *Science* **284**, 92–96.
- Bailey, J. E. (1998) *Biotechnol. Prog.* **14**, 8–20.
- Kueser, H. & Burns, J. A. (1973) *Symp. Soc. Exp. Biol.* **27**, 65–104.
- Fell, D. (1996) *Understanding the Control of Metabolism* (Portland, London).
- Varma, A. & Palsson, B. O. (1994) *Bio/Technology* **12**, 994–998.
- Edwards, J. & Palsson, B. (1999) *J. Biol. Chem.* **274**, 17410–17416.
- Bonarius, H. P. J., Schmid, G. & Tramper, J. (1997) *Trends Biotechnol.* **15**, 308–314.
- Schilling, C. H., Schuster, S., Palsson, B. O. & Heinrich, R. (1999) *Biotechnol. Prog.* **15**, 296–303.
- Liao, J. C., Hou, S. Y. & Chao, Y. P. (1999) *Biotechnol. Bioeng.* **52**, 129–140.
- Schuster, S., Dandekar, T. & Fell, D. A. (1999) *Trends Biotechnol.* **17**, 53–60.
- Mavroulakis, M. & Stephanopoulos, G. (1992) *Comput. Chem. Eng.* **16**, 605–619.
- Kompala, D. S., Ramkrishna, D., Jansen, N. B. & Tsao, G. T. (1986) *Biotechnol. Bioeng.* **28**, 1044–1056.
- Savageau, M. A. (1969) *J. Theor. Biol.* **25**, 365–369.
- Palsson, B. O., Joshi, A. & Ostruz, S. (1987) *Dev. Proc.* **46**, 2485–2489.
- Shu, J. & Shuler, M. G. (1994) *Biotechnol. Bioeng.* **33**, 117–124.
- Lee, I.-D. & Palsson, B. O. (1991) *Biomed. Biochem. Acta* **49**, 771–789.
- Tomita, M., Hashimoto, K., Takahashi, K., Shimizu, T. S., Matsuzaki, Y., Miyoshi, F., Saito, K., Tanida, S., Yugi, K., Venter, J. C., et al. (1999) *Bioinformatics* **15**, 72–84.
- Edwards, J. S., Ramakrishna, R., Schilling, C. H. & Palsson, B. O. (1999) in *Metabolic Engineering*, eds Lee, S. Y. & Papadimitrakaki, E. T. (Dekker, New York), pp. 13–57.
- Sauer, U., Cameron, D. C. & Bailey, J. E. (1998) *Biotechnol. Bioeng.* **59**, 227–238.
- Neidhardt, F. C., ed. (1996) *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Am. Soc. Microbiol., Washington, DC).
- Blattner, F. R., Plunkett, G., Jr., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., et al. (1997) *Science* **277**, 1453–1474.
- Karp, P. D., Riley, M., Saier, M., Paulsen, L. T., Paley, S. M. & Pellegrini-Toole, A. (2000) *Nucleic Acids Res.* **28**, 56–59.
- Selkov, E. Jr., Grechkin, Y., Mikhailova, N. & Selkov, E. (1998) *Nucleic Acids Res.* **26**, 43–45.
- Opata, H., Goto, S., Fujibuchi, W. & Kanehisa, M. (1998) *BioSystems* **47**, 119–128.
- Prismak, J. & Keasling, J. D. (1997) *Biotechnol. Bioeng.* **56**, 398–421.
- Bonarius, H. P. J., Hatzimanikita, V., Meesters, K. P. H., DeGooijer, C. D., Schmid, G. & Tramper, J. (1996) *Biotechnol. Bioeng.* **50**, 299–318.
- Varma, A. & Palsson, B. O. (1993) *J. Theor. Biol.* **165**, 503–522.
- Neidhardt, F. C. & Umberger, H. E. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 13–16.
- Hanson, R. L. & Row, C. (1980) *J. Bacteriol.* **141**, 401–404.
- Frankel, D. G. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 189–198.
- Hutchison, C. A., Peterson, S. N., Gill, S. R., Cline, R. T., White, O., Fraser, C. M., Smith, H. O. & Venter, J. C. (1999) *Science* **286**, 2165–2169.
- Winkler, E. A., Shoemaker, D. D., Astronomoff, A., Ling, H., Anderson, K., Aude, B., Baughman, R., Beiler, R., Beck, J. D., Besser, H., et al. (1999) *Science* **286**, 901–906.
- Edwards, J. S. (1999) Ph.D. thesis (Univ. of California–San Diego, La Jolla).
- Varma, A. & Palsson, B. O. (1994) *Appl. Environ. Microbiol.* **60**, 3724–3731.
- Vanhogen, R. A., Ashkre, K. Z., Moldover, B., Olson, E. R. & Neidhardt, F. C. (1997) *Electrophoresis* **18**, 1253–1251.
- Link, A. J., Robinson, K. & Church, G. M. (1997) *Electrophoresis* **18**, 1259–1313.
- Richmond, C. S., Glasner, J. D., Mau, R., Jin, H. & Blattner, F. R. (1999) *Nucleic Acids Res.* **27**, 3821–3833.
- Brown, P. O. & Botstein, D. (1999) *Nat. Genet.* **21**, 33–37.
- Link, A. J., Phillips, D. & Church, G. M. (1997) *J. Bacteriol.* **179**, 6228–6237.
- Reizer, J., Reizer, A. & Saier, M. H., Jr. (1997) *Microbiology* **143**, 2519–2520.
- Jensen, P. R. & Michelsen, O. (1992) *J. Bacteriol.* **174**, 7635–7641.
- Reitzer, L. J. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 391–407.
- Greene, R. C. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 562–560.
- McFall, E. & Newman, E. B. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 358–379.
- Berlyn, M. K. B., Low, K. B., Rudd, K. E. & Singer, M. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 2, pp. 1715–1902.
- Glasserford, N. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 408–433.
- Matthews, R. G. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 600–611.
- Neuhard, J. & Kell, R. A. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 589–599.
- Penfold, T. & Foster, J. W. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 721–730.
- Jackowski, S. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 681–686.
- White, R. L. & Spenser, I. D. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 680–686.
- Meganathan, R. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 642–656.
- Funk, C. R., Zimlik, & Dowhan, W. (1992) *J. Bacteriol.* **174**, 205–213.
- Mengin-Lecreux, D. & Van Heijenoort, J. (1996) *J. Biol. Chem.* **271**, 32–39.
- Raez, C. R. H. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 1035–1063.
- Postma, P. W., Lengeler, J. W. & Jacobson, G. R. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 1149–1174.
- Lin, E. C. C. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 307–342.
- Cronan, J. E., Jr. & Laporte, D. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 189–198.
- Qin, H., Bongarts, J., Viad, D. & Uden, G. (1997) *Eur. J. Biochem.* **244**, 155–160.
- Craghan, I. T. & Guest, J. R. (1978) *J. Gen. Microbiol.* **107**, 1–13.
- Kumari, S., Tislet, R., Eisenbach, M. & Wolfe, A. J. (1995) *J. Bacteriol.* **177**, 2878–2886.
- Calhoun, M. W., Oden, K. L., Gennis, R. B., de Mattos, M. J. & Neijssel, O. M. (1993) *J. Bacteriol.* **175**, 3020–3025.
- Courtright, J. B. & Hamling, U. (1970) *J. Bacteriol.* **102**, 722–728.
- Vinopal, R. T. & Frankel, D. G. (1974) *J. Bacteriol.* **118**, 1090–1100.
- Singer, M., Walter, W. A., Cali, B. M., Rouviere, P., Liehke, H. H., Gourse, R. L. & Gross, C. A. (1991) *J. Bacteriol.* **173**, 6249–6257.
- Harold, F. M. & Maloney, P. C. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 283–306.
- von Meyeren, B., Jørgensen, B. B., Nielsen, J. & Hansen, F. G. (1982) *Mol. Gen. Genet.* **180**, 248–248.
- Booger, F. C., Boe, L., Michelsen, O. & Jensen, P. R. (1998) *J. Bacteriol.* **180**, 5855–5859.
- Karp, P. D., Krummenacker, M., Paley, S. & Wagg, J. (1999) *Trends Biotechnol.* **17**, 275–281.

**letters to nature**

MDM2. H1299 cells were a gift from J. Chend and were grown in RPMI. All transfections were carried out using Lipofectamine 2000 (Invitrogen). Oligofectamine (Invitrogen) or Genepor2 (Gene Therapy Systems) according to manufacturer's recommendations. To assess the effect of COP1 on steady-state levels of p53, Saos-2 cells were transfected with 100 ng of pcDNA3.1+/-p53, 100 ng of pcMV-Flag-COP1 or pcDNA3.1+/-p53 and U2-Os cells were transfected with 100 ng of pcDNA3.1+/-p53 and 100 ng of pcMV-Flag-COP1 or pcMV-Flag-COP1 and treated with 50.0  $\mu$ M ALLN for 6 h before cell collection where indicated. For reporter assays, Saos-2 or H1299 cells were transiently transfected with 150 ng of pcDNA3.1+/-p53 or pcDNA3.1+/-p53R175H, 100 ng of pcMV-Flag-COP1 or pcMV-Flag-COP1R175H, and 100 ng of pGL3-Luciferase or pGL3-Luciferase with or without increasing amounts (0.5, 1 and 2  $\mu$ g) of pcMV-Flag-COP1 or pcMV-Flag-COP1R175H. Luciferase assays were carried out according to manufacturer's instructions (Promega). For p53-induced cell-death assays, Saos-2 cells were transiently transfected with 1  $\mu$ g of enhanced green fluorescent protein (EGFP) and 5  $\mu$ g of pcDNA3.1+/-p53, pcDNA3.1+/-p53, pcMV-Flag-COP1 or pcDNA3.1+/-p53, 100 ng of pcMV-Flag-COP1 or pcMV-Flag-COP1R175H, and 100 ng of pGL3-Luciferase or pGL3-Luciferase. For analysis by fluorescence-activated cell sorting (FACS), COP1 siRNA (AAGGACCGAA GAUAAUCCUGA), COP1 siRNA1 (AAGAUGUCCUAAUAGAAACCAACGUG), COP1 siRNA2 (AAGACUGGAGGACGAGUUGAUU), COP1 siRNA3 (AAGAGGUGUGUGAG UUGACG), Pirh2 siRNA1 (AACTGTGGATUATTTGAGG), Pirh2 siRNA2 (AACTGTGGATUATTTGAGG), Pirh2 siRNA3 (AACTGTGGATUATTTGAGG), COP1 siRNA4 (AAGGACCGAA GAUAAUCCUGA), COP1 siRNA5 (AAGGACCGAA GAUAAUCCUGA), Pirh2 siRNA4 (AAAUUGUCCGUAUUAUAGG) and MDM2 (AAGGAUUAUUGACCAACCGUAA) siRNA oligonucleotides with 3' dTdT overhangs were synthesized by Genetech or Dharmacon. Control siRNA in experiments refers to a mixture of inverted siRNA oligonucleotides. U2-Os, H1299, Saos-2 and BJ cells were transfected with 100 ng of siRNA three times at 24–36 h intervals and expanded in the presence of 10% FBS. For cell death induction, cells were transfected with 100 ng of siRNA and treated with 100  $\mu$ M etoposide for 48 h.

#### Immunoprecipitation, GST pull-down assays and pulse-chase analysis

Cells were lysed in immunoprecipitation (IP) lysis buffer (1% Triton X-100, 150 mM NaCl, 50 mM Tris, pH 7.4, and protease inhibitor mix) or radioimmunoprecipitation assay (RIPA) buffer (0.1% SDS, 1% NP-40, 150 mM NaCl, 0.5% deoxycholate, 50 mM Tris, pH 7.4, and protease inhibitor mix). Cell lysates were immunoprecipitated with anti-body and protein A/G PLUS beads. Identification of COP1-interacting proteins was carried out as previously described<sup>11</sup>, except that U2-OS cells stably expressing FLAG-COP1 were generated. GST pull-down assays were carried out with GST or GST-SP1-GFP<sup>11</sup> as bait. GST pull-down assays were carried out with GST or GST-SP1-GFP<sup>11</sup> as bait. For 1 h, GST-bound proteins were subject to SDS-PAGE and immunoblotted with anti-HA and anti-GST. IPs were washed in lysis buffer with high salt as required. Pulse-chase experiments were carried out as previously described<sup>11</sup>, except that HEK293T cells were used. Cells were transfected with the indicated plasmids for 24 h, and the cells were transfected with siRNA oligonucleotides as indicated.

### *In vitro* ubiquitination assays

For *in vitro* ubiquitination reactions, *in vitro*-translated p53 was immunoprecipitated with anti-p53 (DO-1 and FL-39) and washed five times with lysis buffer. IP lysate buffer, and reactions were carried out on protein A/G beads. Ten micrograms of Flag-ubiquitin (Sigma), 20 ng of UbcH5b (A.G. Scientific), 50 ng of rabbit E1 (Sigma) and 500 ng of GST-COP1 (E3), which was pre-incubated with 20  $\mu$ M ZnCl<sub>2</sub> for 30 min at room temperature, were incubated in a buffer containing 50 mM Tris, pH 7.5, 2 mM ATP, 5 mM MgCl<sub>2</sub>, 20  $\mu$ M ZnCl<sub>2</sub> and 2 mM DTT. After incubation for 2 h at 30°C with gentle agitation, reactions were stopped by adding 100  $\mu$ M EDTA for 5 min and reduced SDS-PAGE for DSBT for re-immunoprecipitation with anti-p53 (DO-1 and FL-39). Finally, samples were subjected to SDS-PAGE followed by immunoblotting with anti-Flag-FRP (M2) to detect ubiquitinated species of p53.

Received 23 December 2003; accepted 29 March 2004; doi:10.1038/nature02514.  
Published online 21 April 2004.

- MDM2, H1299 cells were a gift from J. Chen. Cells were grown in RPMI all trans-1,25-dihydroxyvitamin D<sub>3</sub> (1,25-(OH)<sub>2</sub>D<sub>3</sub>) promoted the rapid degradation of p53. *Nature* 397, 296–297 (1999).
12. Ling, R. et al. p53, a p53-induced ubiquitin-protein ligase, promotes p53 degradation. *Cell* 112, 779–791 (2003).
13. Thut, C. J., Goodrich, J. A. & Tijun, R. Repression of p53-mediated transcription by MDM2: a dual mechanism. *Genes Dev.* 11, 1574–1586 (1997).
14. Mirrezaei, A. et al. Hdm1 recruits a hypoxia-sensitive complex to negatively regulate p53-dependent transcription. *Curr. Biol.* 13, 1234–1239 (2003).
15. Allan, L. A. & Fried, M. p53-dependent apoptosis or growth arrest induced by different forms of radiation in U2OS cells: p21<sup>WAF1/CIP1</sup> repression in UV induced apoptosis. *Oncogene* 18, 5403–5412 (1999).
16. Qian, H., Wang, T., Naumovski, L., Lopez, C. D. & Bruchman, R. K. Groups of p53 target genes involved in specific p53 downstream effects cluster into different classes of DNA binding sites. *Oncogene* 21, 7901–7911 (2001).
17. Doonan, J. & Hunt, T. Cell cycle? Why don't plants get cancer? *Nature* 380, 481–482 (1996).
18. Wertz, I. E. et al. Human E-cadherin-1 regulates  $\beta$ -catenin by assembling a CUL1A ubiquitin ligase. *Science* 283, 1371–1374 (2000).
19. Doman, D., Shimizu, H., Perkins, N. D. & Hupp, T. R. DNA-dependent acetylation of p53 by the transcription coactivator p300. *J. Biol. Chem.* 274, 13431–13441 (2003).
20. Doman, D. & Hupp, T. R. Inhibition of p53-dependent transcription by BOKA-phospho-peptide mimetics that bind to p300. *EMBO Rep.* 2, 139–144 (2001).
21. Shimizu, H. et al. The conformationally flexible SH-S10 linker region in the core domain of p53 contains a novel MDM2 binding site whose mutation increases ubiquitination of p53 in vivo. *J. Biol. Chem.* 277, 38446–38458 (2002).
- Supplementary Information** The authors provide the paper on [www.nature.com/nature](http://www.nature.com/nature).
- Acknowledgements** We would like to thank T. Hupp for his contribution of reagents and thought-provoking discussions. G. Lozano for p53<sup>+/+</sup>/MDM2<sup>-/-</sup> MEFs, S. Benichou for the Ph12 antibody and cDNA, J. Chen for H1299 cells, B. Henzel for mass spectrometry support, C. Reed for the generation of COP1 monoclonal antibodies, M. Vasser and T. Ng for oligonucleotide synthesis and support for Bioinformatics support, Genentech, Pharmacia Engineering and core DNA sequencing facility for support services, and P. Polakis and members of the Dixit lab for advice and encouragement. I.W. is supported by a PSTF fellowship from the University of California, Davis.
- Competing interests statement** The authors declare that they have no competing financial interests.
- Correspondence** and requests for materials should be addressed to V.M.D. ([cdixit@genecomp.com](mailto:cdixit@genecomp.com)).
- Integrating high-throughput and computational data elucidates bacterial networks**
- Markus W. Covert<sup>1</sup>, Eric M. Knight, Jennifer L. Reed, Markus J. Herrgard & Bernhard O. Palsson
- <sup>1</sup>Biological Engineering Department, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0412, USA
- \* Present address: Biology Division, California Institute of Technology, 1200 E. California Boulevard, Mail Code 147-75, Pasadena, California 91125, USA
- The flood of high-throughput biological data has led to the expectation that computational (or *in silico*) models can be used to direct biological discovery, enabling biologists to reconcile heterogeneous data types, find inconsistencies and systematically generate hypotheses<sup>1–3</sup>. Such a process is fundamentally iterative, where each iteration involves making model predictions, obtaining experimental data, reconciling the predicted outcomes with experimental ones, and using discrepancies to update the *in silico* model. Here we have reconstructed, on the basis of information derived from literature and databases, the first integrated genome-scale computational model of a transcriptional regulatory and metabolic network. The model accounts for 1,010 genes in *Escherichia coli*, including 104 regulatory genes whose products together with other small molecules regulate the expression of 479 of the 906 genes in the recon.

Supplementary Information accompanies the paper on [www.nature.com/nature](http://www.nature.com/nature)

**Acknowledgements** We would like to thank T. Hupp for his contribution of reagents and thought-provoking discussions, G. Lozano for p53<sup>-/-</sup>/MDM2<sup>-/-</sup> MEFs, S. Benichou for the Pirh2 antibody and cDNA, J. Chen for HI299 cells, B. Henzel for mass spectrometry support, C. Reed for the generation of CYP1 monoclonal antibodies, M. Vasser and D. Ng for oligonucleotide synthesis and purification, A. Waugh for Bioinformatics support, Genentech Protein Engineering and core DNA sequencing facility for support services, and P. Polakis and members of the Dixit lab for advice and encouragement. I.W. is supported by a PSTP fellowship from the University of California, Davis.

**Competing interests statement** The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to V.M.D. (dixit@sernec.com).

## Integrating high-throughput and computational data elucidates bacterial networks

Markus W. Covert\*, Eric M. Knight, Jennifer L. Reed,  
Markus J. Herrgard & Bernhard O. Palsson

Bioengineering Department, University of California, San Diego,  
9500 Gilman Drive, La Jolla, California 92093-0412, USA

\* Present address: Biology Division, California Institute of Technology, 1200 E. California Boulevard, Mail Code 147-75, Pasadena, California 91125, USA

The flood of high-throughput biological data has led to the expectation that computational (or *in silico*) models can be used to direct biological discovery, enabling biologists to reconcile heterogeneous data types, find inconsistencies and systematically generate hypotheses<sup>1-3</sup>. Such a process is fundamentally iterative, where each iteration involves making model predictions, obtaining experimental data, reconciling the predicted outcomes with experimental ones, and using discrepancies to update the *in silico* model. Here we have reconstructed, on the basis of information derived from literature and databases, the first integrated genome-scale computational model of a transcriptional regulatory and metabolic network. The model accounts for 1,010 genes in *Escherichia coli*, including 1044 regulatory genes whose products together with other stimuli co-regulate the expression of 479 of the 906 genes in the recon-

structured metabolic network. This model is able not only to predict the outcomes of high-throughput growth phenotyping and gene expression experiments, but also to indicate knowledge gaps and identify previously unknown components and interactions in the regulatory and metabolic networks. We find that a systems biology approach that combines genome-scale experi-

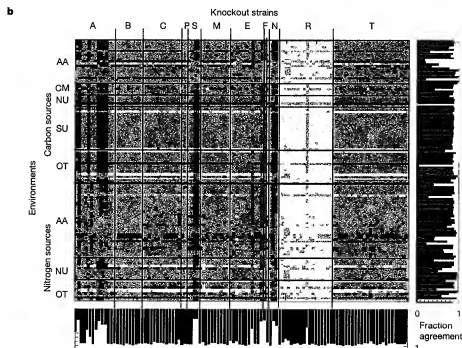
mentation and computation can systematically generate hypotheses on the basis of disparate data sources.

We first validated the model, or 'in silico strain' of *E. coli* (iMC1010<sup>14</sup>; see ref. 4 for conventions for naming in silico strains), against a data set of 13,750 growth phenotypes<sup>5</sup> obtained from the ASAP database<sup>6</sup>, and then used this genome-scale model to select

a	Exp/Met/Reg	Percentage
■	Reg and Met models predict correctly	
	+/+	6,222 45.3%
	-/-	2,094 15.2%
■	Reg model only predicts correctly	
	+/-	657 4.8%
	0	0%
	+/+	1,350 9.8%
	-/-	505 3.7%
■	Met model only predicts correctly	
	+/-	242 1.8%
	-/+	153 1.1%
	0	0%
■	Neither model predicts correctly	
	+/-	702 5.1%
	+/-	1,568 11.4%
■	Met	8,968 65.2%
	Reg	10,828 78.7%
	Total	13,750 100%

c	Carbon sources	FA		Nitrogen sources	FA	
●	Acetoacetic acid	0.42	●	Adenine	0.42	
●	Formic acid	0.05	○	Alanine-Leucine	0.28	
○	γ-Amino butyric acid	0.21	○	Guanosine	0.51	
○	Glycine	0.51	○	L-Lysine	0.44	
○	L-Arginine	0.35	○	L-Methionine	0.42	
○	L-Glutamic acid	0.30	○	L-Phenylalanine	0.24	
○	L-Ornithine	0.21	●	N-Acetyl-D-mannosamine	0.48	
○	Putrescine	0.41	●	Putrescine	0.60	
○	Thymidine	0.00	○	Xanthine	0.04	
Knockout strains			FA	Knockout strains		FA
ΔargB			0.39	○	ΔilvY	0.45
ΔargC			0.31	○	ΔmetA	0.39
ΔargD			0.27	○	Δpgi	0.30
ΔargE			0.52	○	Δpls	0.28
ΔargG			0.57	○	ΔpurD	0.30
ΔglpA			0.25	○	ΔpurH	0.41
ΔglpC			0.25	○	ΔtpiA	0.50
ΔtdvD			0.52	○		

- Missing regulatory interactions
- Missing metabolic enzymes
- Conflicts with subsequent observations



**Figure 1** Growth phenotype study. **a**, Comparison of high-throughput phenotyping array data (Exp) with predictions for the *E. coli* network, both considering regulatory constraints (Reg) and ignoring such constraints as a control (Met). Each case is categorized by comparison type (Exp/Met/Reg), and results are listed as '+' (predicted or observed growth), '-' (no growth) or 'n' (for cases involving a regulatory gene knockout not predictable by the Met model). The comparisons are further divided into four subgroups represented by different colours: AA, amino acid or derivative; CM, central metabolic intermediate; NU, nucleotide or nucleoside; SU, sugar; OT, other. The knockout strains are also divided by

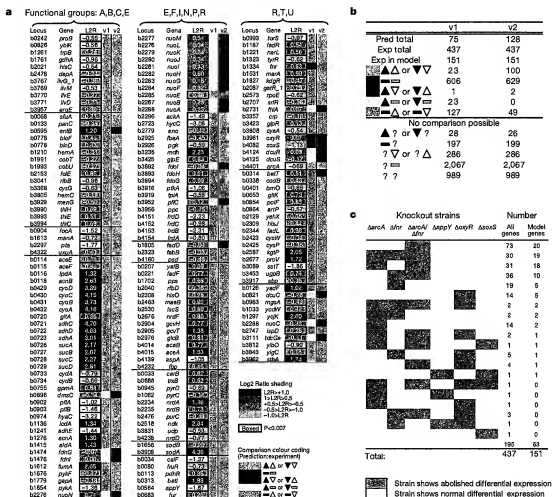
functional group: A, amino acid biosynthesis and metabolism; B, biosynthesis of cofactors, prosthetic groups and carriers; C, carbon compound catabolism; P, cell processes (including adaptation and protection); S, cell structure; M, central intermediary metabolism; E, energy metabolism; F, fatty acid and phospholipid metabolism; N, nucleotide biosynthesis and metabolism; R, regulatory function; T, transport and binding proteins; U, unassigned. Each environment and knockout strain is associated with a fraction of agreement (FA) between regulatory model predictions and observed phenotypes, as shown in the bar charts to the right and below. **c**, Table showing all environments or knockout strains for which FA < 0.60. Of these substrates or knockout strains, 18 point to uncharacterized metabolic or regulatory capabilities in this organism, as indicated (see Supplementary Information for a case-by-case analysis).

transcription factors for prospective gene knockout studies. Comparison with the growth phenotypes showed that experimental and computational outcomes agreed in 10,828 (78.7%) of the cases examined, which is roughly the same success rate achieved in previous studies in *E. coli* and yeast that considered only a few hundred phenotypes<sup>7-9</sup>. In addition, 2,512 (18.3%) of the cases were predicted correctly only when regulatory effects were incorporated into the metabolic model (see Supplementary Information for details).

The comparisons in this study identified several substrates and knockout strains whose growth behaviour did not match predictions (Fig. 1). Further investigation of these conditions and strains led to the identification of five environmental conditions in which dominant, as yet uncharacterized, regulatory interactions actively contribute to the observed growth phenotype, and five environ-

mental conditions and eight knockout strains that highlight uncharacterized enzymes or non-canonical pathways that are predicted to be used by the organism (Fig. 1; a detailed analysis of the discrepancies is provided in the Supplementary Information).

We wanted to determine the utility of this model-driven approach in elucidating transcriptional regulatory networks. A previous study, which evaluated the consistency between existing gene expression data sets and the known transcriptional regulatory network of *E. coli*, identified the response to oxygen deprivation as a partially consistent module<sup>10,11</sup>. We therefore targeted this part of the transcriptional regulatory network for further network characterization. Six strains with knockouts of key transcriptional regulators in the oxygen response ( $\Delta arcA$ ,  $\Delta appY$ ,  $\Delta fnr$ ,  $\Delta oxyR$ ,  $\Delta soxS$  and the double knockout  $\Delta arcA\Delta fnr$ ) were constructed. The messenger RNA expression profiles of these strains, as well as the



wild-type strain, were measured in aerobic and anaerobic glucose minimal medium conditions. The data were analysed<sup>12</sup> in the context of iMC1010<sup>11</sup> predictions to identify new interactions in the regulatory network (Fig. 2).

Expression profiling of the wild-type strain identified 437 genes that experienced a significant change in transcription in response to oxygen deprivation (*t*-test, multiple testing corrected to give a false discovery rate (FDR) of less than 5%); of these, 151 genes were included in iMC1010<sup>11</sup>. Computationally, 75 genes were predicted by iMC1010<sup>11</sup> to show differential expression in response to oxygen deprivation. These 75 genes could be divided into three categories: 23 agreed with measured expression changes; 24 had a predicted expression change that was either not found to be statistically significant in the experimental data (23/24) or in a direction opposite to that of the experimental data (1/24); and for 28 genes there were no expression data available (transcript abundance was determined to be 'absent' for two or more of the replicates). Thus, of the 47 (=23 + 24) differentially expressed genes that could be compared between the model computation and experiment, 23 (or 49% accuracy) agreed. Considering the overall number of genes in the model for which there were experimental data, the overlap (23) between the sets of predicted (47) and experimentally detected (151) differentially expressed genes is significant in comparison to a model that would randomly predict expression changes ( $P < 0.005$  on the basis of a cumulative binomial distribution). There were 151 genes that were differentially expressed and included in the model; however, with only 23 (or 15% coverage) correctly computed, there is much room for expanding the transcriptional regulatory network in iMC1010<sup>11</sup> on the basis of the experimental data (Fig. 3).

To understand which transcription factors are involved in regulating these differentially expressed genes after oxygen deprivation, we compared the gene expression data for the wild-type and each knockout strain separately. Using two-way analysis of variance

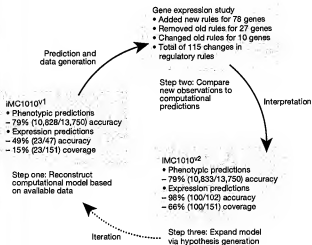
(ANOVA), we could determine whether the differential expression was significantly altered in the knockout strain as compared with the wild type. A large portion of the expression changes observed for the wild-type strain were not significantly affected in any of the knockout strains (195/437 or 44.6% of genes overall, 63/151 or 41.7% of genes in the model, FDR < 5%), suggesting that none of the five transcription factors studied here regulates the expression of these genes or that combinatorial interactions between multiple transcription factors are involved in regulation. For the remainder of the genes, differential expression was abolished in one or more of the knockout strains (Fig. 2c).

The ANOVA-based identification of transcription factors that influence differential expression of specific genes enabled us systematically to rewrite, relax or remove various regulatory rules in the model to resolve the discrepancies between iMC1010<sup>11</sup> and the experimentally determined wild-type differential gene expression. For many (81) of the genes, a regulatory rule already existed and had to be reconciled with our new data to accommodate the newly determined transcription factor dependencies. For genes where none of the knockouts abolished differential expression, we simply based a new regulatory rule on the presence of oxygen rather than a transcription factor (39 genes). By contrast, for genes where a change in expression was predicted but not observed, we removed oxygen dependency from the existing regulatory rule (23 genes). In addition, for 12 genes the predicted expression changes agreed with the observed expression in the wild type, but our knockout perturbation analysis indicated that the transcription factors involved in the regulation differed from previously reported data and the model needed to be changed (all new regulatory rules are detailed in the Supplementary Information).

The updated model (iMC1010<sup>12</sup>) was used to recalculate all of the predictions for both the aerobic and anaerobic expression data and the high-throughput phenotyping arrays. Note that iMC1010<sup>12</sup> accounts for the same genes as iMC1010<sup>11</sup> but has different regulatory interactions among the gene products and oxygen as an environmental variable. We found agreement between model predictions and the gene expression data to be substantially higher using the iMC1010<sup>12</sup> model, as expected (Fig. 2c). Specifically, 100 of the 151 expression changes were correctly computed with iMC1010<sup>12</sup>, and the number of false-positive predictions (Fig. 2, yellow boxes) was reduced to zero. In resolving many of the cases of unpredicted differential expression (Fig. 2, orange boxes), we found that implementation of the ANOVA-derived rule resulted in the inability of the wild-type or knockout *in silico* strain to grow aerobically or anaerobically on glucose, or under other conditions where growth had been previously established (for example, wild-type and knockout strain average growth rate under aerobic conditions,  $0.68 \pm 0.04$  per hour; anaerobic,  $0.43 \pm 0.07$  per hour). Such cases may be thought of as an 'overfit' of the microarray data. Accordingly, we relaxed the regulatory rule for these genes (42 in total) to allow for a correct phenotype prediction. Comparisons for the high-throughput phenotyping data revealed very little difference from Fig. 1 (only 11 out of 13,750 cases were affected; see Supplementary Information).

The iterative modification of the regulatory rules led to three main observations. First, some of the results of the knockout perturbation analysis are complex enough to make boolean rule formulation difficult. For example, the interplay of Fnr and ArcA can lead to complex behaviours where the expression change observed in wild type is abolished in the  $\Delta arcA$  or the  $\Delta fnr$  strains, but not in the  $\Delta arcA \Delta fnr$  strain. Such complex interplay among the transcription factors can lead to specialized expression changes, as observed in the *gdyAB* response to anaerobic, microaerobic and aerobic conditions<sup>13,14</sup>.

Second, in revising regulatory rules for transcription factors we found that whereas in some cases, such as *arcA*, expression of a regulatory protein correlates positively with its activity, in several



**Figure 3** Biological network elucidation by a model-centric approach. Metabolic and regulatory networks may be expanded by using high-throughput phenotyping and gene expression data coupled with the predictions of a computational model. If model predictions are consistent with experimental observations, the network is adequately characterized. If not, the model identifies a knowledge gap and may be used to update, validate and generate hypotheses about organism function. Accuracy refers to the percentage of model predictions that agree with experimental data; coverage indicates the percentage of experimental changes predicted correctly by the model.

cases, including *fnr*, *btf* and *fur* among others, the mRNA level of a regulatory gene is reduced when the protein is in fact activated. For example, under anaerobic conditions when *Fnr* is known to be active<sup>15</sup>, its expression is significantly reduced. Such behaviour, underscored by similar observations of mRNA transcript levels and corresponding protein product abundance in yeast<sup>16</sup>, suggests that the identification of regulatory networks, and transcription factors involved in regulation in particular, will not be accomplished by the determination of co-regulated gene sets alone.

Third, many of these gene expression changes involve complex interactions and indirect effects. Transcription factors may be affected, for example, by the presence of fermentation by-products or the build up of internal metabolites. Such effects would be extremely difficult to identify or account for without a computational model.

In summary, we find that the reconciliation of high-throughput data sets with genome-scale computational model predictions enables systematic and effective identification of new components and interactions in microbial biological networks. Our study illustrates only the first round of an iterative model building strategy where an initial model based on literature-derived information (MC1010<sup>11</sup>) is used to design informative experiments and then updated on the basis of the new experimental data obtained (MC1010<sup>12</sup>). Another round of perturbation experiments will lead to MC1010<sup>13</sup>, and so on. We expect that after an effort of some years and many iterations of this process, regulatory network elucidation for *E. coli* will be essentially complete. □

## Methods

### Computational model

We constructed the model of the *E. coli* metabolic and regulatory network by identifying network components, their functions and interactions from the primary literature<sup>17,18,19</sup>. Many approaches have been developed to analyse large-scale metabolic<sup>20–22</sup> and transcriptional regulatory<sup>23–25</sup> networks. Growth and gene expression simulations were done by regulated flux-balance analysis, which combines linear optimization to determine a growth-optimized metabolic flux distribution with logic statements to simulate the effects of regulatory processes over time. The whole model construction and simulation process has been described elsewhere in detail<sup>18</sup>.

### Strains and culture

The parent strain for knockout strains in this study was K-12 MG1655 (ref. 27), and all deletion strains were generated as described<sup>27</sup>. Growth experiments for the gene expression study were done on M9 glucose medium (12 g l<sup>-1</sup>) under aerobic and anaerobic conditions, as described<sup>27</sup>. The growth data contained in the ASAP database were obtained by using high-throughput phenotype arrays (Biolog<sup>®</sup>). In some cases (where the viability of a particular environment was unclear from the phenotype array data), we cross-validated the ASAP phenotyping data by culturing the wild-type strain under the given conditions in our laboratory (see Supplementary Information).

### Gene expression profiling and analysis

All gene expression measurements were done at least in triplicate. Samples were stabilized by using RNeasy Protect Bacterial reagent (Qiagen), and total RNA was isolated from exponentially growing cells using a RNeasy mini kit (Qiagen) in accordance with the manufacturer's protocols (see <http://www.qiagen.com>). The RNA (10 µg) was then used as the template for complementary DNA synthesis, the product of which was fragmented, labelled and hybridized to an *E. coli* Antisense Gene Array (Affymetrix), which was washed and scanned to obtain an image in accordance with the manufacturer's protocols (see <http://www.affymetrix.com>). The image files were processed and expression values were normalized using dChip software<sup>28</sup>. We used quantitative real-time polymerase chain reaction with reverse transcription (RT-PCR) to validate expression changes for selected genes. The statistical significance of expression changes for each gene and each strain between aerobic and anaerobic conditions was determined by a *t*-test (log-transformed data, equal variance).

For each deletion strain, we used a two-way ANOVA (strain as the first factor and aerobic or anaerobic condition as the second factor) to determine whether the differential expression observed in the wild-type strain was significantly altered in the deletion strain by determining the statistical significance of the strain-condition interaction effect. For both the *t*-test and the ANOVA analysis, correction for multiple testing was done by using the Benjamini-Hochberg false discovery rate procedure<sup>29</sup>, which determines the *P*-value cut-off for each test separately by estimating the FDR resulting from using a particular *P*-value cut-off. The false discovery rate refers to the fraction of true null tests out of all the tests called significant and an FDR of 5% was used for all tests. All gene expression data and

the relevant information (such as the MIAME checklist) are provided in the Supplementary Information.

Received 22 November 2003; accepted 1 March 2004; doi:10.1038/nature02456.

1. Ideker, T., Galitski, T. & Hood, L. A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet.* 2, 343–372 (2001).
2. Palsson, B. O. The challenges of *in silico* biology. *Nature Biotechnol.* 18, 1147–1150 (2000).
3. Kitano, H. Systems biology: a brief overview. *Science* 295, 1662–1664 (2002).
4. Reed, J. L., Vo, T. D., Schilling, C. H. & Palsson, B. O. An expanded genome-scale model of *Escherichia coli* K-12 (BW04 GSM/GPR). *Genome Biol.* 4, R54–R54 (2003).
5. Buchner, B. R. New technologies to assess genotype-phenotype relationships. *Nature Rev. Genet.* 4, 309–314 (2003).
6. Glaser, J. D. et al. ASAP: a systematic annotation package for community analysis of genomes. *Nucleic Acids Res.* 31, 1471–1481 (2003).
7. Foster, J., Familli, J., Palsson, B. O. & Nielsen, J. Large-scale evaluation of *in silico* gene knockouts in *Saccharomyces cerevisiae*. *Omicron* 7, 193–202 (2003).
8. Edwards, J. S. & Palsson, B. O. The *Escherichia coli* MG1655 *in silico* metabolic genotype: its definition, characteristics, and capabilities. *Proc. Natl Acad. Sci. USA* 97, 5528–5533 (2000).
9. Covert, M. W. & Palsson, B. O. Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*. *J. Biol. Chem.* 277, 18009–18014 (2002).
10. Herrgard, M. J., Covert, M. W. & Palsson, B. O. Reconciling gene expression data with known genome-scale regulatory network structures. *Genome Res.* 13, 2423–2434 (2003).
11. Salamon, K. et al. Global gene expression profiling in *Escherichia coli* K12. The effects of oxygen availability and FNR. *J. Biol. Chem.* 278, 29837–29855 (2003).
12. Ideker, T. et al. Integrated genomic and proteomic analysis of a systematically perturbed metabolic network. *Science* 292, 929–934 (2001).
13. Compain, J. & Tsoukalas, D. Anaerobic activation of *arcA* transcription in *Escherichia coli*: roles of *Fur* and *ArcA*. *Mol. Microbiol.* 11, 955–964 (1994).
14. Coffee, P. A., Melville, S. B., Albrecht, J. A. & Gunsalus, R. P. Aerobic regulation of cytochrome *d* oxidase (*cydAB*) operon expression in *Escherichia coli*: roles of *Fur* and *ArcA* in repression and activation. *Mol. Microbiol.* 25, 605–615 (1997).
15. Griffin, T. J. et al. Comprehensive profiling of gene expression at the transcriptome and proteome levels in *Saccharomyces cerevisiae*. *Mol. Cell Proteomics* 1, 323–333 (2002).
16. Reed, J. L. & Palsson, B. O. Thirteen years of building constraint-based models of *Escherichia coli*. *J. Bacteriol.* 185, 2892–2899 (2003).
17. Edwards, J. S., Ihara, R. U. & Palsson, B. O. *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nature Biotechnol.* 19, 125–130 (2001).
18. Price, N. D., Papin, J. A., Schilling, C. H. & Palsson, B. O. Genome-scale microbial *in silico* models: the constraints-based approach. *Trends Biotechnol.* 21, 162–169 (2003).
19. Segre, D., Vitkup, D. & Church, G. M. Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl Acad. Sci. USA* 99, 15112–15117 (2002).
20. Burgard, A. P. & Moras, C. D. Probing the performance limits of the *Escherichia coli* metabolic network subject to gene deletions or deletions. *Biochem. Biophys. Res. Commun.* 294, 364–375 (2001).
21. Burgard, A. P., Vaidyanathan, S. & Moras, C. D. Minimal reaction sets for *Escherichia coli* metabolism under different growth requirements and uptake environments. *Biochem. Biophys. Res. Commun.* 279, 791–799 (2001).
22. Jeong, H., Tomb, B., Albert, R., Oltvai, Z. N. & Barabási, A. L. The large-scale organization of metabolic networks. *Nature* 407, 651–654 (2000).
23. Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genet.* 31, 64–68 (2002).
24. Gutierrez-Rios, R. M. et al. Regulatory network of *Escherichia coli*: consistency between literature knowledge and microarray profiles. *Genome Res.* 13, 2435–2443 (2003).
25. Ben-Joseph, Z. et al. Computational discovery of gene modules and regulatory networks. *Nature Biotechnol.* 21, 1337–1342 (2003).
26. Covert, M. W., Schilling, C. H. & Palsson, B. Regulation of gene expression in flux balance models of metabolism. *J. Theor. Biol.* 213, 73–88 (2001).
27. Blattner, F. R. et al. The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453–1474 (1997).
28. Datar, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA* 97, 6640–6645 (2000).
29. Li, C. & Wong, W. H. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl Acad. Sci. USA* 96, 31–36 (1999).
30. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300 (1995).

Supplementary information accompanies the paper on [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank K. Sadokovic and A. Fleming for assistance with computation; Z. Zhang and A. Raghunathan for experimental assistance; the Perna and Blattner laboratories for access to the high-throughput phenotyping data in the ASAP database; and the NIH for funding and support. M.W.C. and B.O.P. designed the project and were involved in all phases of the study; E.M.C. carried out experiments; J.L.R. reconstructed the model, ran simulations and did the phenotyping analysis; M.J.H. did the statistical analysis of the gene expression data.

**Competing interests statement** The authors declare competing financial interests: details accompany the paper on [www.nature.com/nature](http://www.nature.com/nature)

**Correspondence** and requests for materials should be addressed to B.O.P. (palsson@ucsd.edu). The gene expression data are available online in GEO (<http://www.ncbi.nlm.nih.gov/geo/>), accession number GSE1121.

## Phagocytosis and conidiocidal assays

Alveolar macrophages were exposed to conidia at a macrophage:conidia ratio of 1:5 for 2 h, and 1:1 for 4 h, in the absence or presence of 20 µg ml<sup>-1</sup> FITC (0.44 × 10<sup>6</sup> M FITC protomoy) before being evaluated for internalization or conidiocidal activity, respectively (Supplementary Information).

Received 26 July; accepted 8 October 2002; doi:10.1038/nature01195.

1. Emley, J. et al. Structure of pentameric human serum amyloid P component. *Nature* **367**, 338–345 (1994).
2. Szalai, A. J., Agrawal, A., Greenough, T. J. & Volanakis, J. E. C-reactive protein structural biology, gene expression, and host defense function. *Immunol. Rev.* **16**, 127–136 (1997).
3. Steel, D. M. & Whitehead, A. S. The major acute phase reactant C-reactive protein: serum amyloid P component and serum amyloid A protein. *Immunol. Today* **15**, 81–88 (1994).
4. Papp, M. B. & Balazs, M. L. Acute phase proteins with special reference to C-reactive protein and related proteins (pentaxins) and serum amyloid A protein. *Adv. Immunol.* **34**, 141–212 (1983).
5. Nurnsterg, M. et al. Role of serum amyloid P component in bacterial infection: protection of the host or protection of the pathogen. *Proc. Natl Acad. Sci. USA* **97**, 14584–14589 (2000).
6. Mantovani, A., Garlanda, C. & Bottazzi, B. *Cytokine Reference* (eds Oppenheim, J. & Feldmann, M.) (Academic, New York, 2002).
7. Breviario, S. et al. Interleukin-1-inducible genes in endothelial cells. Cloning of a new gene related to C-reactive protein and serum amyloid P component. *J. Biol. Chem.* **267**, 22190–22197 (1992).
8. Lee, G. W., Lee, T. H. & Vilek, J. TSG-14, a tumour necrosis factor- and IL-1-inducible protein, is a novel member of the pentaxin family of acute phase proteins. *J. Immunol.* **150**, 1804–1812 (1993).
9. Vidali Allos, V. et al. Inducible expression of PTX3, a new member of the pentaxin family, in human mononuclear phagocytes. *Blood* **84**, 3483–3493 (1994).
10. Introna, M. et al. Cloning of mouse PTX3, a new member of the pentaxin gene family expressed at extrathymic sites. *Blood* **87**, 1862–1872 (1996).
11. Bottazzi, B. et al. Multimer formation and ligand recognition by the long pentraxin PTX3 – similarities and differences with the short pentraxin C-reactive protein and serum amyloid P component. *J. Biol. Chem.* **272**, 32817–32822 (1997).
12. Rovere, P. et al. The long pentraxin PTX3 binds to apoptotic cells and regulates their clearance by antigen-presenting dendritic cells. *Blood* **96**, 4300–4306 (2000).
13. Lee, G. W., Goodman, A. R., Lee, T. H. & Vilek, J. Relationship of TSG-14 protein to the pentaxin family of major acute phase proteins. *J. Immunol.* **153**, 3700–3707 (1994).
14. Vazani, S. et al. Knockout of pentaxin-3, a downstream target of growth differentiation factor-9, causes female subfertility. *Mol. Endocrinol.* **16**, 1154–1167 (2002).
15. Medfordin, B. Toll-like receptors and innate immunity. *Nature Rev. Immunol.* **1**, 135–145 (2001).
16. Denning, D. W. Invasive aspergillosis. *Clin. Infect. Dis.* **26**, 781–803 (1998).
17. Cenci, E. et al. Th1 and Th2 cytokines in mice with invasive aspergillosis. *Infect. Immun.* **65**, 564–570 (1997).
18. Menacaci, A. et al. Defective interleukin-12-helper 1 (Th1) immunity in a murine model of allogeneic T-cell-depleted bone marrow transplantation and its restoration by treatment with Th2 cytokine antagonists. *Blood* **97**, 1483–1490 (2001).
19. Schleifner, A., Douglas, H. & Brude, A. Selective protection against conidia by mononuclear and against mycelia by polymorphonuclear phagocytes in resistance to *Aspergillus*. Observations on these two lines of defense in vivo and in vitro with human and mouse phagocytes. *J. Clin. Invest.* **69**, 617–631 (1982).
20. Ramani, L. The T cell response against fungal infections. *Curr. Opin. Immunol.* **9**, 484–490 (1997).
21. Menacaci, A. et al. Cytokines in candidiasis and aspergillosis. *Curr. Pharm. Biotechnol.* **1**, 235–251 (2000).
22. Bozza, S. et al. Dendritic cells transport conidia and hyphae of *Aspergillus fumigatus* from the airways to the draining lymph nodes and initiate disparate Th responses to the fungus. *J. Immunol.* **168**, 1363–1371 (2002).
23. Botto, M. et al. Homozygous C1q deficiency causes glomerulonephritis associated with multiple apoptotic bodies. *Nature Genet.* **19**, 56–59 (1998).
24. Peri, G. et al. PTX3, a prototypic long pentraxin, is an early indicator of acute myocardial infarction in man. *Circulation* **102**, 636–641 (2000).
25. Muller, B. et al. Circulating levels of the long pentraxin PTX3 correlate with severity of infection in critically ill patients. *Crit. Care Med.* **29**, 1406–1407 (2001).
26. Villa, D. et al. Pattern of cytokines and chemokines in sepsis induced by cecal ligation and puncture compared with that induced by endotoxin. *Clin. Diagn. Lab. Immunol.* **2**, 549–553 (1995).
27. Watanabe, Y., Miyazawa, M., Sano, M., Nakano, H. & Nomoto, K. Enhanced resistance against *Listeria monocytogenes* in an early phase of primary infection in pregnant mice: activation of macrophages during pregnancy. *Infect. Immun.* **52**, 730–735 (1984).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

**Acknowledgements** This work was supported by Istituto Superiore di Sanità, Ministero Istruzione, Università e Ricerca (MIUR), Consiglio Nazionale delle Ricerche (CNR), and by the European Commission. We acknowledge the contribution of the Italian Association for Cancer Research. We thank C. Scotton for critical reading of the manuscript.

**Competing interests statement** The authors declare competing financial interests: details accompany the paper on Nature's website (<http://www.nature.com/nature>).

**Correspondence** and requests for materials should be addressed to A.M. (e-mail: mantovani@marionegri.it).

## *Escherichia coli* K-12 undergoes adaptive evolution to achieve *in silico* predicted optimal growth

Rafael U. Ibarra\*, Jeremy S. Edwards†‡ & Bernhard O. Palsson\*

\* Department of Bioengineering, University of California, San Diego,

9500 Gilman Drive, La Jolla, California 92093-0412, USA

‡ Department of Chemical Engineering, University of Delaware, Newark, Delaware 19716, USA

† These authors contributed equally to this work

Annotated genome sequences<sup>1,2</sup> can be used to reconstruct whole-cell metabolic networks<sup>3–6</sup>. These metabolic networks can be modelled and analysed (computed) to study complex biological functions<sup>7–11</sup>. In particular, constraints-based *in silico* models<sup>12</sup> have been used to calculate optimal growth rates on common carbon substrates, and the results were found to be consistent with experimental data under many but not all conditions<sup>13,14</sup>. Optimal biological functions are acquired through an evolutionary process. Thus, incorrect predictions of *in silico* models based on optimal performance criteria may be due to incomplete adaptive evolution under the conditions examined. *Escherichia coli* K-12 MG1655 grows sub-optimally on glycerol as the sole carbon source. Here we show that when placed under growth selection pressure, the growth rate of *E. coli* on glycerol reproducibly evolved over 40 days, or about 700 generations, from a sub-optimal value to the optimal growth rate predicted from a whole-cell *in silico* model. These results open the possibility of using adaptive evolution of entire metabolic networks to realize metabolic states that have been determined a priori based on *in silico* analysis.

Predictive whole-cell metabolic models can be developed using constraints-based modelling procedure<sup>15–18</sup>. As an alternative to detailed theory-based models, constraints-based models use the successive imposition of governing constraints (such as mass conservation, thermodynamics, capacity and nutritional environment) to eliminate network functions that exceed the governing constraints. Mathematically this procedure defines a solution space containing all possible metabolic network functions that satisfy the governing constraints. Each particular solution in this space corresponds to a particular state of the metabolic network and therefore a potential behaviour of the cell. Within the solution space defined by the governing constraints, the optimal use of the metabolic network to support growth can be found among all possible solutions using linear optimization<sup>16–19</sup>. However, a single optimal growth condition is of limited interest and a phenotype phase plane (PPP) analysis has been developed to obtain a broad understanding of a metabolic network's optimal properties<sup>20,21</sup>. The PPP analysis evaluates the optimal properties of a metabolic network under a range of environmental conditions (see Methods) and has been used to show that the growth of *E. coli* is consistent with the optimal use of its metabolic network under several defined growth conditions<sup>13–14</sup>.

It is not known whether optimal growth is observed on all substrates, and if not, whether adaptive evolution towards optimal growth can be achieved. Furthermore, if such adaptive evolution towards the optimal behaviour occurs, does the endpoint correspond with a priori calculations? To address these issues, we examined prolonged exponential growth of *E. coli* K-12 on several substrates (acetate, succinate, malate, glucose and glycerol). All calculations presented here were made with a previously formulated large-scale *E. coli* metabolic model<sup>13,14</sup>, and the model was not adjusted or 'fitted' to the data described.

Batch growth experiments were done using malate as the sole



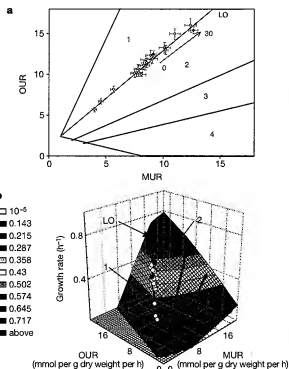
carbon source with a range of substrate concentrations ( $0.25\text{--}3\text{ g l}^{-1}$ ) and temperatures ( $29\text{--}37^\circ\text{C}$ ) to vary the malate uptake rate (MUR). The MUR, oxygen uptake rate (OUR) and growth rate were measured. The measured MUR and OUR data were optimal, as defined by the line of optimality (LO) in the PPP (Fig. 1a). The optimal growth rate of *E. coli* was calculated for all combinations of the MUR and OUR and displayed as a surface over the PPP (Fig. 1b). The experimentally determined growth rates were on the edge of the colour-coded solution space that corresponds to the LO (Fig. 1b). Hence the optimal growth performance of *E. coli* K-12 on malate was predicted *a priori* by using PPP. The results for growth with malate as the sole carbon source were in agreement with previous observations of *E. coli* metabolism for growth on succinate or acetate<sup>14</sup>.

A natural question arises: is the optimal performance on malate stable over prolonged periods of time? To address this question, adaptive evolution of *E. coli* on malate was studied for 500 generations. The adaptation resulted in a 19% increase in growth rate. However, the MUR and OUR also increased and maintained metabolic operation on the LO (Fig. 1). Similar adaptive evolution experiments on acetate and succinate resulted in an increased growth rate (20% and 17%, respectively) (Fig. 2). Both the oxygen and substrate uptake rates increased concomitantly to maintain optimal growth as defined and predicted by the PPP analysis.

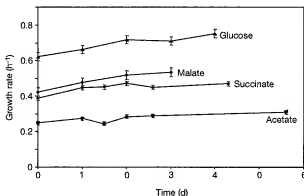
The growth rate of *E. coli* using glucose as the carbon source was also increased by prolonged exponential growth (Figs 2 and 3). Before adaptive evolution on glucose the cellular growth rate, OUR, and glucose uptake rate (GUR) were experimentally determined over a range of glucose concentrations and temperatures. The experimentally determined values for the GUR and OUR corresponded to points on the LO or slightly in phase 2 (the acetate overflow region) of the PPP (ref. 21) (Fig. 3a). The predicted acetate secretion in phase 2 was experimentally observed and the measured growth rates were on the surface of the solution space near the edge corresponding to the LO (data not shown). *E. coli* was subsequently kept in sustained exponential growth over 500 generations (Fig. 3b, c). The growth rate increased by 17%, as shown by movement of the experimental data points within phase 2 on the surface towards higher growth rates. Thus, as with malate, succinate and acetate, the growth rate of *E. coli* with glucose as the carbon source could be slightly increased with the substrate and oxygen uptake rates moving in phase 2 with some acetate overflow. It was also noted that evolutionary adaptation maintained metabolic operation on the surface of the three-dimensional PPP, as predicted by the physicochemical constraints on the metabolic network. The metabolic operation in the phase 2 provided an increased growth rate

with a reduced yield (relative to the LO).

We determined the growth performance over a range of glycerol concentrations and temperature. Unlike growth on malate or glucose, the experimental data points were scattered throughout phase 1 far from the LO (Fig. 4b), indicating sub-optimal growth of



**Figure 1** Growth of *E. coli* K-12 on malate. **a**, The malate-oxygen phenotype phase plane (PPP). Phase 1 is characterized by metabolic futile cycles, whereas phase 2 is characterized by acetate overflow metabolism. The line of optimality (LO, in red) separates phases 1 and 2 (ref. 21). Data points (open circles) represent malate concentrations ranging from  $0.25\text{--}3\text{ g l}^{-1}$ , and temperatures ranging from  $29\text{--}37^\circ\text{C}$ . The two data points in blue represent the starting point (day 0) and endpoint (day 30) of adaptive evolution respectively, at a malate concentration of  $2\text{ g l}^{-1}$  and a temperature of  $37^\circ\text{C}$ . These data points represent a span of 500 generations. **b**, Three-dimensional representation of growth rates. The x and y axes represent the same variables as in **a**. The z axis represents the cellular growth rate ( $\text{h}^{-1}$ ). OUR, oxygen uptake rate; MUR, malate uptake rate.



**Figure 2** Growth rate during adaptive evolution on glucose, malate, succinate and acetate. Growth conditions were kept constant at a temperature of  $37^\circ\text{C}$  and a substrate concentration of  $2\text{ g l}^{-1}$ . We measured growth rate in the exponential phase of growth.

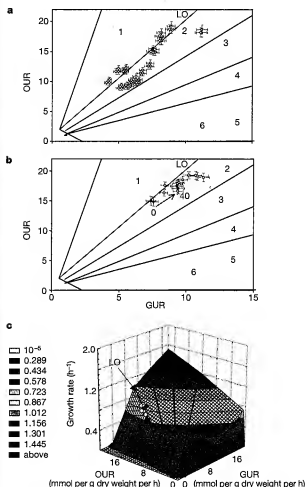
The increases in growth rate over time were as follows: glucose (18%), malate (21%), succinate (17%) and acetate (20%). The number of generations for each adaptive evolution was: glucose (500), malate (500), succinate (1,000) and acetate (700).

wild-type *E. coli* K-12 on glycerol, consistent with previous observations<sup>22</sup>.

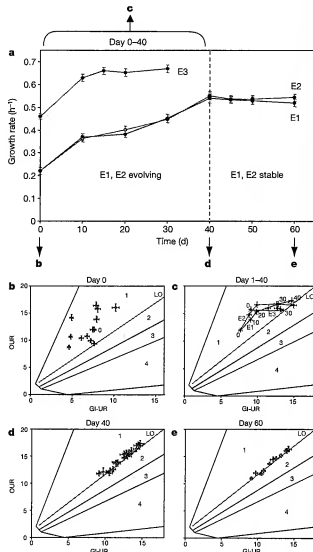
We studied *E. coli* adaptive growth using glycerol as the sole carbon source ( $2 \text{ g l}^{-1}$ ), by serial transfer, at a temperature of  $30^\circ\text{C}$ , and with sufficient oxygenation. Growth rate, glycerol uptake rate (G1-UR) and OUR were measured every ten days. Over a 40-day period an evolutionary path (E1) was observed (Fig. 4c). Phenotypic changes were traced in phase 1, eventually converging towards the LO. During this 40-day period, the growth rate more than doubled from  $0.23 \text{ h}^{-1}$  to  $0.55 \text{ h}^{-1}$  (Fig. 4a). The substrate uptake and growth rate data obtained under various growth conditions after adaptive evolution were near the LO (Fig. 4d). The evolved strain attained near-optimal growth on glycerol as defined by the *in silico* predictions. A second, independent adaptation experiment gave a similar but non-identical evolutionary trajectory (E2), converging near the same endpoint (Fig. 4c). Finally, a third independent adaptation experiment (E3) was done with a different initial starting point within phase 1. E3 was done at  $37^\circ\text{C}$  and a glycerol concentration of  $2 \text{ g l}^{-1}$ . The adaptation of *E. coli* to growth

on glycerol at  $37^\circ\text{C}$  resulted in motion towards the LO and the growth rate increased by about 30% (Fig. 4a). The final growth rate of the E3 strain was consistent with the *in silico* predictions with respect to the G1-UR, OUR and the growth rate.

To assess the stability of the endpoint of the adaptive evolution, we extended the cultivation on glycerol for an additional 300 generations, or 20 days for the E1 and E2 strains. The data indicated no further change in growth (Fig. 4a). On the sixtieth day of the experiment the E1 and E2 strains exhibited growth on the LO under various growth conditions, reaffirming optimal growth behaviour and the stability of the phenotype (Fig. 4c).



**Figure 3** Growth of *E. coli* K-12 on glucose. **a**, The glucose-oxygen PPP. Like the malate-oxygen PPP, phase 1 represents sub-optimal growth and phase 2 is characterized by acetate overflow metabolism. The LO is shown in red. **b**, The PPP plotted against OUR along with experimental values for adaptive evolutionary experiments. Open circles represent measurements during the adaptive evolutionary process, whereas blue circles indicate the beginning and end of evolution. **c**, Three-dimensional rendering of computed growth rate and the experimental data (from **a** and **b**). The  $x$  and  $y$  axes represent the GUR and OUR. The  $z$  axis represents the cellular growth rate ( $\text{h}^{-1}$ ).



**Figure 4** Growth of *E. coli* K-12 on glycerol. **a**, Change in growth rate with time for three adaptive evolution experiments: trajectories E1, E2, and E3. E1 and E2 were performed at  $30^\circ\text{C}$  and E3 at  $37^\circ\text{C}$ . The glycerol concentration was kept constant at  $2 \text{ g l}^{-1}$  during E1, E2 and E3. **b**, The PPP pre-evolution. The LO is shown in red. The range of glycerol concentrations used was  $0.25\text{--}2 \text{ g l}^{-1}$ . **c**, The PPP during adaptive evolution. Experimental values for E1 are indicated in blue, and for E2 they are indicated in green, and for E3 in red. The starting point of evolution for E1 and E2 is indicated in black (day 0). **d**, The PPP after 40 days (about 700 generations) of evolution. The range of glycerol concentrations used was  $0.25\text{--}2 \text{ g l}^{-1}$ . **e**, The PPP after 60 days (1,000 generations) of evolution. The range of glycerol concentrations used was  $0.25\text{--}2 \text{ g l}^{-1}$ . Data points were obtained using the E1 (blue) and E2 (green) strains. G1-UR, glycerol uptake rate.

Selection pressure is expected to result in optimal performance during an evolutionary process. Optimal growth of *E. coli* on acetate, succinate, malate and glucose is consistent with the predictions of whole-cell *in silico* models. The strain used here has presumably never had to compete for survival using glycerol as the sole carbon source and thus initially utilized this carbon source non-optimally. However, adaptive evolution on glycerol resulted in the *a priori* calculated optimal growth that was based on the constraints placed on the *E. coli* metabolic network. The adaptive evolutionary process had a reproducible and predictable endpoint.

This study opens up several possibilities. First, it may now be possible to specify optimal network properties *in silico* and achieve them through an adaptive evolutionary process or in combination with a series of other methodologies. *In silico* design of microorganisms could be used to improve their metabolic abilities, production efficiency and/or operational longevity. Second, changes in mRNA expression levels and DNA sequences can now be monitored as cells progress along a defined evolutionary path. Such experiments may yield valuable insight into the molecular design of complex control circuits and their adaptation during evolution. The combination of *in silico* and experimental biology introduced here may make a new series of biological designs attainable.

Constraint-based computational models use an optimization-based procedure to predict cellular states. It is assumed that this optimal state, within the governing constraints, is found by altering the numerical values of the kinetic and regulatory constants through a 'trial-and-error' process. This feature of constraint-based models is a significant departure from other types of mathematical models of cell function, where these parameters are treated as being time-invariant. Thus constraint-based models directly account for the fundamental nature of adaptive evolution. The adaptive evolutionary path itself cannot be predicted; however, the final outcome can be. □

## Methods

### Strains and media

The *E. coli* K-12 MG1655 annotated genome sequence and the biochemical literature were used to construct the *in silico* *E. coli* strain<sup>10,11</sup>. We simulated the metabolic capabilities as previously described with the objective of maximizing growth<sup>12,13,14</sup>. Both growth and maintenance requirements were imposed on the *in silico* model<sup>15,16</sup>. The growth experiments were done in M9 minimal medium with the addition of the carbon source. The growth rate was varied by changing the concentration of the carbon source (ranging from 0.25 to 3 g l<sup>-1</sup>) and the temperature (ranging from 29°C to 37°C). The *MG1655* (ATCC 47078) was used for all of the experiments.

Batch cultures were set up at two different volume scales. One-litre (large) cultures were performed in 1.5-l Erlenmeyer flasks sparging with air. The large-volume batch cultures were used to continuously monitor the oxygen uptake rate (OUR) online with an off-gas analyser. Small (100–250 ml) cultures were grown in 500–1,000 ml Erlenmeyer flasks. For the small-scale cultures the OUR was monitored online polarographically and by measuring the mass transfer coefficient for oxygen (*k<sub>La</sub>*) (see below). The temperature was controlled by using a circulating water bath (Haake). We measured and analysed data during exponential growth. The biomass and the concentration of the substrate (malate, glucose, glycerol) in the media were monitored throughout the experiment.

### Analytical procedures

Cellular growth rate was monitored by measuring the absorbance (*A*, or optical density) at 600 nm and 420 nm and by cell counts (Coulter Electronics). The doubling time was calculated from the growth rate:  $t_d = \ln(2)/\mu$ . Absorbance to cellular dry weight correlations were determined by two different measurements: (1) spun-down cells were dried at 75°C to a constant weight; and (2) 25–50 ml samples (taken throughout the culture) were filtered, washed and dried to a constant weight. The concentration of metabolites in the culture media was determined by high-performance liquid chromatography (HPLC) (Rainin Instruments). An aminex HPX-87H ion exchange carbohydrate-organic acid column (Bio-Rad Laboratories) (65°C) was used with degassed 5 mM sulphuric acid as the mobile phase and ultraviolet detection. Glucose and glycerol were monitored by enzymatic assay (Sigma). The dissolved oxygen in the culture was monitored with a polarographic dissolved oxygen probe (Cole-Parmer Instruments). Oxygen consumption was measured in three different ways: (1) passing the effluent gas through a Servomex oxygen analyser (model 1140C) (Servomex); (2) calculated from the dissolved oxygen reading and *k<sub>La</sub>* measurements; and (3) in a respirometer chamber in a separate 70-ml flask. All three methods used for measuring the OUR gave similar and reproducible results.

### Adaptive evolution

Cultures in prolonged exponential growth were started from individual colonies and were grown in micro-carrier spinner flasks at 250 ml within a temperature-controlled incubator. Serial transfers were made during the exponential phase of growth at mid-log-phase ( $A_{600nm} = 0.55$ ) using an adjusted inoculum volume based on the growth rate of the culture. On a daily basis, the growth rate, time of inoculation,  $A_{600nm}$  of the culture, and any visual changes in the composition of the culture were recorded. Samples of cultures were stored on a daily basis. The culture was tested weekly for pH and phenotyped (by plating) for any signs of coexistence with a distinct population of mutants or foreign contamination. No discernible differences in colony morphology or signs of foreign contamination were observed during these experiments.

### Phenotype phase plane analysis

First, the metabolic reconstruction was done using biochemistry, genomic and physiological data<sup>17,18</sup>. Second, mass balance, capacity and thermodynamic constraints were imposed on the network to define a solution space<sup>19</sup>. Third, the best use of the metabolic network for a given objective was identified using linear optimization<sup>16–18</sup>. Fourth, all optimal solutions as a function of two constraints were presented on the PPP<sup>14</sup>. The PPP has distinct phases. Each phase corresponds to a particular type of an optimal solution, which in turn represents a particular flux distribution through the network. Each phase has certain metabolic characteristics; for instance, phase 2 in the figure presented here was characterized by an acetate overflow. The lines that demarcate the phases were defined by changes in the shadow price structure of the optimal solution. The properties of the PPP have been detailed elsewhere<sup>11</sup>. In this study, the optimal utilization of the metabolic network was predicted *a priori*, based on a previously developed *in silico* model<sup>12,14</sup>.

Received 12 December 2001; accepted 2 September 2002; doi:10.1038/nature01149.

- Blattner, F. R. et al. The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1474 (1997).
- Drell, D. The Department of Energy Microbial Cell Project: a 180° paradigm shift for biology. *OMICS* **6**, 3–9 (2002).
- Covett, M. W. et al. Metabolic modeling of microbial strains *in silico*. *Trends Biochem. Sci.* **26**, 179–186 (2001).
- Sellou, E., Maltsev, N., Olsen, G. J., Overbeck, R. & Whitman, W. B. A reconstruction of the metabolism of *Methanococcus jannaschii* from sequence data. *Gene* **197**, GC11–GC26 (1997).
- Overbeck, R. et al. WTE integrated system for high-throughput genome sequence analysis and metabolic reconstruction. *Nucleic Acids Res.* **28**, 123–125 (2000).
- Kang, P. D. et al. The Ecolink database. *Nucleic Acids Res.* **30**, 56–58 (2002).
- Gombert, A. K. & Nielsen, J. Mathematical modeling of metabolism. *Curr. Opin. Biotechnol.* **11**, 180–186 (2000).
- Tomioka, M. et al. E-CELL: software environment for whole-cell simulation. *Bioinformatics* **15**, 72–84 (1999).
- Voll, D. *Understanding the Control of Metabolism* (Portland, London, 1996).
- Schuster, S., Fell, D. A. & Dandekar, T. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nature Biotechnol.* **18**, 326–332 (2000).
- Schilling, C. H. et al. Genome-scale metabolic model of *Helicobacter pylori* 26695. *J. Bacteriol.* **184**, 4582–4593 (2002).
- Edwards, J. S. & Palsson, B. O. The *Escherichia coli* MG1655 *in silico* metabolic genome: its definition, characteristics, and capabilities. *Proc. Natl. Acad. Sci. USA* **97**, 5528–5533 (2000).
- Verma, A. & Palsson, B. O. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl. Environ. Microbiol.* **60**, 3724–3731 (1994).
- Edwards, J. S., Barrar, R. U. & Palsson, B. O. *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nature Biotechnol.* **19**, 125–130 (2001).
- Palsson, B. O. The challenges of *in silico* biology. *Nature Biotechnol.* **18**, 1147–1150 (2000).
- Edwards, J. S., Ramakrishna, R., Schilling, C. H. & Palsson, B. O. *Metabolic Engineering* (eds Lee, S. Y. & Papadimitrakis, F. T.) (Wiley, New York, 1999).
- Bonarius, H. P. J., Schmitt, O. & Tramper, J. Flux analysis of undetermined metabolic networks: The quest for the missing constraints. *Trends Biotechnol.* **15**, 308–314 (1997).
- Verma, A. & Palsson, B. O. Metabolic flux balancing: basic concepts, scientific and practical use. *BioTechnology* **12**, 994–998 (1994).
- Wiechert, W. Modeling and simulation tools for metabolic engineering. *J. Biotechnol.* **94**, 37–53 (2002).
- Schilling, C. H., Edwards, J. S., Letecher, D. & Palsson, B. O. Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems. *Biochem. Biophys. Res. Commun.* **271**, 286–306 (2000).
- Edwards, J. S., Ramakrishna, R. & Palsson, B. O. Characterizing the metabolic phenotype: a phenotype phase plane analysis. *Biotechnol. Bioeng.* **77**, 27–36 (2002).
- Wiley, C. C., Sauer, U. & Bailey, J. E. Use of a glycerol-limited, long-term chemostat for isolation of *Escherichia coli* mutants with improved physiological properties. *Microbiology* **143**, 1567–1574 (1997).

**Acknowledgements** We thank L. Ramon, J. DiToro, S. Fong, J. Macielinski, N. Short and H. Bialy for technical assistance and for reviewing the manuscript. We acknowledge funding support from the National Institutes of Health and the National Science Foundation, and the Department of Energy Office of Biological and Environmental Research.

**Competing interests statement** The authors declare competing financial interests: details accompany the paper on *Nature's website* (http://www.nature.com/nature).

**Correspondence** and requests for materials should be addressed to B.O.P. (e-mail: palsson@ucsd.edu).

Docket No.: 066662-0092

**PATENT**

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re Application of	:	Customer Number: 41552
	:	
Palsson, Bernhard	:	Confirmation Number: 1729
	:	
Application No.: 09/923,870	:	Group Art Unit: 1631
	:	
Filed: August 06, 2001	:	Examiner: Negin, Russell Scott
	:	
For: METHODS FOR IDENTIFYING DRUG TARGETS BASED ON GENOMIC SEQUENCE DATA	:	

**DECLARATION UNDER 37 C.F.R. § 1.132**

Mail Stop Amendment  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

I, Jens B. Nielsen, declare as follows:

1. I am a Professor in Systems Biology at Chalmers University of Technology in Göteborg, Sweden. I previously was a full professor at the Technical University of Denmark and served as both the deputy director and the director of Center for Process Biotechnology. I am founder of the Center for Microbial Biotechnology and served as its director until 2007.

2. I earned a Masters of Science in chemical engineering and a Ph.D in biotechnology from the Technical University of Denmark. I did a post doctoral fellowship at the University of Hannover in Germany. I have authored or co-authored more than 270 papers in international peer review journals and am the co-author of three text books in the field of bioreaction engineering, metabolic engineering and metabolomics. I am an inventor or co-inventor on more than 25 issued patents and I am a founder of four companies. I also am a member of Genomatica's Scientific Advisory Board. I am a recipient of a number of prestigious awards, including the Villum Kann Rasmussen's Årslegat and the Merck Award for Metabolic Engineering. A copy of my curriculum vitae and a list of publications is attached as Exhibit 1.

EXHIBIT C

3. I am very familiar with stoichiometric models of metabolism and have read U.S. application serial no. 09/923,870, by Palsson. I also am very familiar with Dr. Palsson's work, including the publication that is the basis of this application (Edwards and Palsson, *Proc. Natl. Acad. Sci. U.S.A.*, 97:5528-33 (2000)). I understand that the invention described in this application is directed to constructing genome specific stoichiometric matrices that can be utilized with flux balance analysis for modeling metabolism. The application claims, in part, a method of simulating a metabolic capability by incorporating metabolic reactions through the use of genome information. I also understand that the claimed invention stands rejected for obviousness over the combination of references to Pramanik and Keasling, *Biotech. and Bioengineering* 56:398-421 (1997) in view of Blattner et al., *Science* 277:1453-69 (1997) and in view of Kunst et al., *Rev. in Microbiol.* 142:905-12 (1991).

4. I know from personal knowledge that at the time Professor Palsson's invention was made respected and prestigious scientists in the field of metabolic engineering publically criticized Professor Palsson's invention and had a strong disbelief that it worked.

5. In July of 1999 I chaired a meeting session at the Biochemical Engineering XI meeting held in Salt Lake City, Utah. I also was a member of this meeting's Advisory Committee. Attached as Exhibit 2 is a copy of the final program for that meeting entitled Biochemical Engineering XI: Molecular Diversity in Discovery and Bioprocessing.

6. During the meeting session that I chaired, Professor Palsson was giving a lecture on the use of metabolic models for simulations of growth and analysis of deletion mutants and other characteristics using the concepts of flux balance analysis. After Professor Palsson's lecture, Professor Jay Bailey from the auditorium requested whether he could provide some comments to the lecture. I agreed as I assumed that it was the normal type of questions and comments that are asked of lecturers at conferences, and considering Professor Baileys prominent position in the field of biochemical engineering and metabolic engineering I expected a good discussion. In fact Professor Bailey has published himself in the field of metabolic flux analysis, and had been working quite a lot on metabolic models.

7. As it turned out, Professor Bailey had prepared 4-5 overhead slides analyzing and criticizing the approach taken by Professor Palsson. It was clear to me that there must have been

Application No.: 09/923,870

some earlier correspondence, and it also seemed like Professor Bailey had prepared "opposition" to another lecture than the one Professor Palsson had just given, as several of the critical points raised by Professor Bailey were in fact addressed in the lecture of Professor Palsson.

8. The main criticism raised by Professor Bailey was that it was not possible to predict metabolic functions and cellular physiology by simply using constraint based simulations. He argued that due to the large degrees of freedom the predictions are not likely to represent true phenotypes. It is correct that there is a large degrees of freedom in the system and that constraint based simulations cannot predict correctly all the fluxes in the network, but correct phenotypes have been demonstrated in numerous examples later to be captured very well by this kind of simulation. The "opposition" from Professor Bailey was supported by several others in the auditorium, so it was quite clear that there was a general belief that the concept would not work.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements are made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that any such willful false statement may jeopardize the validity of the application or any patent issued thereon.

  
\_\_\_\_\_  
Jens B. Nielsen  
\_\_\_\_\_  
Date

# Curriculum vitae

## Jens Nielsen

Department of Chemical and Biological Engineering, Chalmers University of Technology,  
Kemigården 4, SE-412 96 Gothenburg, Sweden

Tel: +46 (0) 31 772 3804; Fax: +46 (0) 31 772 3801; E-mail: [nielsenj@chalmers.se](mailto:nielsenj@chalmers.se)

### General information

Date of Birth: November 17, 1962 in Horsens, Denmark  
Citizenship: Danish  
Home Address: Nedre Fogelbergsgatan 5, SE-411 28 Gothenburg, Sweden

### Education and degrees

1986 M.Sc. in Chemical Engineering, Dept. of Biotechnology, DTU, Denmark  
1989 Ph.D. in Biochemical Engineering, DTU, Denmark  
1995 dr.techn., DTU, Denmark

### Employments

1986 **Research Assistant**, School of Engineering, University of Western Ontario, Canada  
1987-1989 **PhD student**, Department of Biotechnology, DTU, Denmark  
1989 **Post doc**, Institut für Technische Chemie, Universität Hannover, Germany  
1990-1995 **Associate Research Professor**, Department of Biotechnology, DTU, Denmark  
1995-1996 **Visiting Professor**, Department of Chemical Engineering, MIT, USA  
1996-1998 **Associate Professor**, Department of Biotechnology, DTU, Denmark  
1998-2008 **Professor**, BioCentrum, DTU, Denmark  
2005- **Adjunct Professor**, Department of Biotechnology, NTNU, Norway  
2008- **Professor**, Department of Chemical and Biological Engineering, Chalmers

### Academic Experience

1994-1999 **Founding Chairman**, Danish Biotechnological Forum, Denmark  
1995-2000 **Deputy Director**, Center for Process Biotechnology, DTU, Denmark  
1995-1998 **Co-ordinator**, Physiological Engineering, Nordic Industrial Fund  
1998-2002 **Co-ordinator**, NordPhys, Nordic Industrial Fund  
2001-2003 **Director**, Center for Process Biotechnology, DTU, Denmark  
2001-2003 **Board Member**, Center for BioProcess Technology, KTH, Sweden  
2002-2007 **SAB Member**, Kluwyer Center for Genomics of Industrial Fermentations, TU Delft, The Netherlands  
2002-2008 **SAB Member**, Max-Planck Institute for Dynamic of Complex Technical Systems, Germany  
2004-2007 **Founding Director**, Center for Microbial Biotechnology, DTU, Denmark  
2004-2007 **Member of Integrative and Systems Biology Panel**, BBSRC, UK (50 million GPD)  
2004- **Advisory Board**, Society for Biological Engineering, USA  
2005 **Chairman of Evaluation Committee**, HepatoSys, Germany  
2005-2008 **Co-ordinator**, Yeast Systems Biology Network, EU Framework VI  
2006-2007 **Founder and Member**, Danish Biotechnological Society, Denmark  
2008- **Member of Review Panel of SystemsX.ch**, Switzerland (100 million CHF)  
2008- **Co-ordinator**, Systems Biology as a Driver for Industrial Biotechnology, EU Framework VII

**Society Memberships**

1986-	<b>The Danish Society for Engineers, Denmark</b>
1996-	<b>Member of American Association for the Advancement of Science, USA</b>
1997-	<b>Member of the Academy of Technical Sciences, Denmark</b>
2001-	<b>Society for Industrial Microbiology, USA</b>
2004-	<b>American Chemical Society, USA</b>
2004-	<b>Society for Biological Engineering, USA</b>
2006-	<b>American Society for Microbiology, USA</b>

**Publications, Patents & Presentations**

284	<b>Peer Reviewed Papers</b>
20	<b>Book Contributions</b>
3	<b>Text books:</b> Bioreaction Engineering (1994,2003) (English & Chinese), Metabolic Engineering (1998) (English, Chinese & Japanese), Metabolome Analysis (2007)
1	<b>Monograph</b>
3	<b>Edited books</b>
28	<b>Issue patents (12 patent families)</b>
>95	<b>Invited oral presentations at international conferences</b>
>45	<b>Invited seminars at universities and companies</b>

**Key Grants**

1993-1996	<b>Nordic Project on Physiological Engineering, Nordic Industrial Fund (&gt;20 million DKK)</b>
1995-2002	<b>4 Projects, EU Framework VI (about 8 million DKK)</b>
1996-1999	<b>Center for Process Biotechnology, Danish Technical Research Council (20 million DKK)</b>
1997-2000	<b>Nordic network NordPhys, Nordic Industrial Fund (&gt;4 million DKK)</b>
1999-2003	<b>DABIC, Danish Technical Research Council (18 million DKK)</b>
2000-2007	<b>6 Projects, EU Framework V (about 15 million DKK)</b>
2003-2007	<b>Heterologous production of polyketides, Danish Technical Research Council (2.8 million DKK)</b>
2004-2008	<b>Center for Microbial Biotechnology, Danish Technical Research Council (36 million DKK)</b>
2005-2009	<b>3 Projects, EU Framework VI (about 8.5 million DKK)</b>
2005-2008	<b>Genome sequencing of <i>A. baumannii</i>, Lundbeck Foundation (1.2 million DKK)</b>
2005-2008	<b>Yeast Systems Biology Network, EU Framework VI (1.3 million €)</b>
2006-2008	<b>Yeast in No Gravity, European Space Agency (2.7 million DKK)</b>
2007-2009	<b>Systems Biology: From model organisms to application, NORDFORSK (0.8 million €)</b>
2008-2011	<b>Systems Biology as a Driver for Industrial Biotechnology, EU Framework VII (1 million €)</b>
2008-2013	<b>UNICELLSYS, EU Framework VII (5.4 million SEK)</b>

**Awards**

1989	<b>Direktør Gorm Petersen's Mindelegat, Denmark</b>
1994	<b>Ulrik Brinch og Hustru Marie Brinch's legat, Denmark</b>
1995	<b>Fullbright Fellow, USA</b>
1996	<b>STVFs Jubilæumspris, Denmark</b>
2001	<b>Aksel Tovborg Jensens Legat/Bjerrum-Brøndsted-Lang Lecture, Denmark</b>
2002	<b>Villum Kann Rasmussen's Årslegat, Denmark</b>



- 2002 **Sunner Memorial Lecture**, Lund University, Sweden  
 2004 **Merck Award for Metabolic Engineering**, USA  
 2004 **Hough Memorial Lecture**, Birmingham University, UK

#### Editorial Activities

- 1996 **Chemical Engineering Science**, Guest editor of special volume  
 1999- **Metabolic Engineering**, Editorial Board (1999-2002), Associate Editor (2003-)  
 1999- **Applied Microbiology and Biotechnology**, Editorial Board  
 2000- **Bioprocess and Biosystems Engineering**, Associate Editor (2000-2007), Editorial Board (2007-)  
 2001- **Biotechnology and Bioengineering**, Associate Editor  
 2001- **FEMS Yeast Research**, Editorial Board (2001-2007), Associate Editor (2007-)  
 2004- **Journal of Industrial Microbiology and Biotechnology**, Editorial Board  
 2004- **Advances in Biochemical Engineering/Biotechnology**, Editorial Board

#### Teaching Experience

Has organized more than 15 advanced courses on metabolic engineering and systems biology in Denmark, Sweden, Thailand, China and Chile  
 Has been teaching in several different courses at all levels in the area of biotechnology, fermentation technology, bioreaction engineering, metabolic engineering and systems biology.  
 Have extensive experience with setting up, organizing and running new courses.  
 Have been actively involved in the design of new teaching programs at both the BSc and MSc level in the field of biotechnology and systems biology.

#### Mentoring Experience

Main supervisor of graduated PhD students	43
Co-supervisor of graduated PhD students	22
Current main supervisor of PhD students	12
Current co-supervisor of PhD students	2
Former affiliated post docs	34
Current affiliated post docs	8

#### Business Experience

- 1996- **Consultant** for several internationally leading biotech and pharmaceutical companies  
 1997 **GeneCare ApS**, Denmark, Founder  
 1997-2001 **N&N Biotechn ApS**, Denmark, Founder and CEO  
 2001-2004 **Symbion Venture Capital**, Denmark, Member of Scientific Advisory Board  
 2002- **Fluxome Sciences A/S**, Denmark, Founder, CEO (2002-2004), CSO (2002-2008), Member of BOD, Chairman of SAB (2008-) (has raised more than 7 million € in capital)  
 2002-2007 **Gothia Yeast Solutions**, Sweden, Member of SAB  
 2006- **MycTeQ A/S**, Denmark, Founder and Chairman of BOD (has raised about 1 million € in capital)  
 2008- **GlycoFi, Inc.**, USA, Member of SAB  
 2008- **Promethegen, Inc.**, USA, Member of SAB

#### Organization of Conferences

2000-	Member of organizing and scientific committee of more than 25 conferences
2000	<b>ESBES3</b> , Denmark, Chairman (~300 delegates)
2002	<b>Analysis of Microbial Cells at the Single Cell Level</b> , Denmark, Co-Chair (~100 delegates)
2002	<b>Metabolic Engineering IV</b> , Italy, Chairman (~250 delegates)
2004	<b>European Conference on Fungal Genetics VI</b> , Denmark, Co-Chair (~800 delegates)
2005	<b>ECB12</b> , Denmark, Chairman of Scientific Committee (~1200 delegates)

## List of Publications for Jens Nielsen

### Research papers in international journals with peer review

#### 1988

- (1) K. Nikolajsen; J. Nielsen; J. Villadsen (1988)  
In-line flow injection analysis for monitoring lactic acid fermentations. *Anal. Chim. Acta* **214**:137-145

#### 1989

- (2) J. Nielsen; K. Nikolajsen; J. Villadsen (1989)  
FIA for on-line monitoring of important lactic acid fermentation variables. *Biotechnol. Bioeng.* **33**:1127-1134
- (3) J. Nielsen; C. Emborg; K. Halberg; J. Villadsen (1989)  
Compartment model concept used in the design of fermentation with recombinant microorganisms *Biotechnol. Bioeng.* **34**:478-486

#### 1990

- (4) J. Nielsen; K. Nikolajsen; S. Benthin; J. Villadsen (1990)  
Application of flow-injection analysis in the on-line monitoring of sugars, lactic acid, protein, and biomass during lactic acid fermentations. *Anal. Chim. Acta* **237**:165-175
- (5) G. Wehnert; K.-D. Anders; B. Bittner; R. Kammeyer; U. Hübner; J. Nielsen; T. Scheper (1990)  
Ein kombinierter Fluoreszenz-/Streulichtsensor und dessen Einsatz zur Prozessbeobachtung in der Biotechnologie. *Chem.-Ing.-Tech.* **62**:211-212

#### 1991

- (6) S. Benthin; J. Nielsen; J. Villadsen (1991)  
A simple and reliable method for the determination of cellular RNA content. *Biotechnol. Technol.* **5**:39-42
- (7) J. Nielsen; K. Nikolajsen; J. Villadsen (1991)  
Structured modelling of a microbial system 1. A theoretical study of the lactic acid fermentation. *Biotechnol. Bioeng.* **38**:1-10
- (8) J. Nielsen; K. Nikolajsen; J. Villadsen (1991)  
Structured modelling of a microbial system 2. Experimental verification of a structured lactic acid fermentation model. *Biotechnol. Bioeng.* **38**:11-23
- (9) K. Nikolajsen; J. Nielsen; J. Villadsen (1991)  
Structured modelling of a microbial system 3. Growth on mixed substrates. *Biotechnol. Bioeng.* **38**:24-29
- (10) J. Nielsen; A. G. Pedersen; K. Strudsholm; J. Villadsen (1991)  
Modelling fermentations with recombinant microorganisms : Formulation of a structured model. *Biotechnol. Bioeng.* **37**:802-808
- (11) S. Benthin; J. Nielsen; J. Villadsen (1991)  
Characterisation and application of precise and robust flow injection analyzers for on-line measurement during fermentations. *Anal. Chim. Acta* **247**:45-50
- (12) L. H. Christensen; J. Nielsen; J. Villadsen (1991)  
Monitoring of substrates and products during fed-batch penicillin fermentations on complex media. *Anal. Chim. Acta* **249**:123-136
- (13) L. H. Christensen; J. Nielsen; J. Villadsen (1991)  
Delay and dispersion in an in-situ membrane probe for bioreactors. *Chem. Eng. Sci.* **46**:3304-3307

#### 1992

- (14) S. Benthin; J. Nielsen; J. Villadsen (1992)  
Flow Injection Analysis of micromolar concentrations of glucose and lactate in fermentation media. *Anal. Chim. Acta* **261**:145-153
- (15) S. Benthin; J. Nielsen; J. Villadsen (1992)  
Anomeric specificity of glucose uptake systems in *Lactococcus cremoris*, *Escherichia coli* and *Saccharomyces cerevisiae*: Mechanisms, kinetics and implications. *Biotechnol. Bioeng.* **40**:137-146
- (16) J. Nielsen (1992)  
On-line monitoring of microbial processes by flow injection analysis. *Proc. Control Qual.* **2**:371-384
- (17) C. L. Johansen; L. H. Christensen; J. Villadsen; J. Nielsen (1992)  
Monitoring and control of fed-batch penicillin fermentation. *Comp. Chem. Eng.* **16**:S297-S304
- (18) K. Strudsholm; J. Nielsen; C. Emborg (1992)  
Product formation during batch fermentation with recombinant *Escherichia coli* containing a runaway plasmid. *Bioprocess Eng.* **8**:173-181

#### 1993

- (19) A. G. Pedersen; M. Bundgaard; O. Hassager; J. Nielsen; J. Villadsen (1993)  
Rheological characterization of media containing *Penicillium chrysogenum*. *Biotechnol. Bioeng.* **41**:162-164
- (20) J. Nielsen (1993)  
A simple morphologically structured model describing the growth of filamentous microorganisms. *Biotechnol. Bioeng.* **41**:715-727

- (21) M. Carlsen; L. H. Christensen; J. Nielsen (1993)  
Flow-injection analysis for measurement of penicillin V in fermentation samples. *Anal. Chim. Acta* 274:117-123
  - (22) M. Carlsen; H. Meier; F. Lantrebecq; C. Johansen; R. W. Min; J. Nielsen (1993)  
On-line monitoring of penicillin V during penicillin fermentations : A comparison of two different methods based on FIA. *Anal. Chim. Acta* 279:51-58
  - (23) S. Benthin; J. Nielsen; J. Villadsen (1993)  
Transport of sugars via two anomer-specific sites on mannose-phosphotransferase system in *Lactococcus cremoris* : In vivo study of mechanism, kinetics and adaption. *Biotechnol. Bioeng.* 42:440-448
  - (24) J. Nielsen (1993)  
Simulation of bioreactions. *Comp. Chem. Eng.* 18:S615-S620
  - (25) S. Benthin; J. Nielsen; J. Villadsen (1993)  
Two uptake systems for fructose in *Lactococcus lactis* subsp. *cremoris* FD1 produce glucolytic and gluconeogenic fructose phosphates and induce oscillations of growth and lactic acid formation. *Appl. Environ. Microbiol.* 59:3206-3211
- 1994**
- (26) S. Benthin; J. Nielsen; J. Villadsen (1994)  
Galactose expulsion during lactose metabolism in *Lactococcus lactis* subsp. *cremoris* FD1 due to dephosphorylation of intracellular galactose-6-phosphate. *Appl. Environ. Microbiol.* 60:1254-1259
  - (27) A. G. Pedersen; M. Bundgaard-Nielsen; J. Nielsen; J. Villadsen (1994)  
Characterization of mixing in stirred tank bioreactors equipped with Rushton turbines. *Biotechnol. Bioeng.* 44:1013-1017
  - (28) A. G. Pedersen; H. Andersen; J. Nielsen; J. Villadsen (1994)  
A novel technique based on  $^{85}Kr$  for quantification of gas-liquid mass transfer in bioreactors. *Chem. Eng. Sci.* 49:803-810
  - (29) S. Benthin; U. Schulze; J. Nielsen; J. Villadsen (1994)  
Growth energetics of *Lactococcus cremoris* FD1 during energy-, carbon- and nitrogen limitation in steady state and transient cultures. *Chem. Eng. Sci.* 49:589-610
  - (30) L. H. Christensen; G. Mandrup; J. Nielsen; J. Villadsen (1994)  
A robust LC method for studying the penicillin fermentation. *Anal. Chim. Acta* 296:51-62
  - (31) L. H. Christensen; J. Nielsen; J. Villadsen (1994)  
Degradation of penicillin V in fermentation media. *Biotechnol. Bioeng.* 44:165-169
  - (32) M. A. Hjortso; J. Nielsen (1994)  
A conceptual model of autonomous oscillations in microbial cultures. *Chem. Eng. Sci.* 49:1083-1095
  - (33) J. Nielsen; C. L. Johansen; J. Villadsen (1994)  
Culture fluorescence measurements during batch and fed-batch cultivations with *Penicillium chrysogenum*. *J. Biotechnol.* 38:51-62
  - (34) M. Carlsen; J. Marcher; J. Nielsen (1994)  
An improved FIA-system for measuring  $\alpha$ -amylase in cultivation media. *Biotechnol. Tech.* 8:479-482
- 1995**
- (35) H. Jørgensen; J. Nielsen; J. Villadsen; H. Mølgaard (1995)  
Analysis of the penicillin V biosynthesis during fed-batch cultivations with a high yielding strain of *Penicillium chrysogenum*. *Appl. Microbiol. Biotechnol.* 43:123-130
  - (36) J. Nielsen; P. Krabben (1995)  
Hyphal growth and fragmentation of *P. chrysogenum* in submerged cultures. *Biotechnol. Bioeng.* 46:588-598
  - (37) J. Nielsen; C. L. Johansen; M. Jacobsen; P. Krabben; J. Villadsen (1995)  
Pellet formation and fragmentation in submerged cultures of *Penicillium chrysogenum* and its relation to penicillin production. *Biotechnol. Prog.* 11:93-98
  - (38) H. Jørgensen; J. Nielsen; J. Villadsen; H. Mølgaard (1995)  
Metabolic flux distributions in *Penicillium chrysogenum* during fed-batch cultivations. *Biotechnol. Bioeng.* 46:117-131
  - (39) R. Mørkeberg; M. Carlsen; J. Nielsen (1995)  
Induction and repression of  $\alpha$ -amylase production in recombinant and wild-type strains of *Aspergillus oryzae*. *Microbiol.* 141:2449-2454
  - (40) J. Nielsen; H. S. Jørgensen (1995)  
Metabolic control analysis of the penicillin biosynthetic pathway in a high yielding strain of *Penicillium chrysogenum*. *Biotechnol. Prog.* 11:299-305
  - (41) R. Lejeune; J. Nielsen; G. Baron (1995)  
Morphology of *Trichoderma reesei* QM 9414 in submerged cultures. *Biotechnol. Bioeng.* 47:609-615
  - (42) L. H. Christensen; U. Schulze; J. Nielsen; J. Villadsen (1995)  
Acoustic gas analysis for fast and precise monitoring of bioreactors. *Chem. Eng. Sci.* 50:2101-2110
  - (43) R. Lejeune; J. Nielsen; G. Baron (1995)  
Influence of pH on the morphology of *Trichoderma reesei* QM 9414 in submerged cultures. *Biotechnol. Lett.* 17:341-344
  - (44) Rong Wei Min; J. Nielsen; J. Villadsen (1995)

- Simultaneous monitoring of glucose, lactic acid and penicillin by Sequential Injection Analysis. *Anal. Chim. Acta* **312**:149-156
- (45) L. H. Christensen; C. M. Henriksen; J. Nielsen; J. Villadsen; M. Egel-Mitani (1995)  
Continuous cultivation of *P. chrysogenum*. Growth on glucose and penicillin production. *J. Biotechnol.* **42**:95-107
- (46) M. A. Hjortso; J. Nielsen (1995)  
Population balance models of autonomous microbial oscillations. *J. Biotechnol.* **42**:255-269
- (47) Rong Wei Min; M. Carlsen; J. Nielsen; J. Villadsen (1995)  
Measurements of  $\alpha$ -amylase activity by Sequential Injection Analysis. *Biotechnol. Techn.* **9**:765-768
- 1996**
- (48) M. Carlsen; J. Nielsen; J. Villadsen (1996)  
Kinetic studies of acid-inactivation of  $\alpha$ -amylase from *Aspergillus oryzae*. *Chem. Eng. Sci.* **51**:37-43
- (49) M. Carlsen; A. B. Spøhr; J. Nielsen; J. Villadsen (1996)  
Morphology and physiology of an  $\alpha$ -amylase producing strain of *Aspergillus oryzae* during batch cultivations. *Biotechnol. Bioeng.* **49**:266-276
- (50) C. M. Henriksen; L. H. Christensen; J. Nielsen; J. Villadsen (1996)  
Growth energetics and metabolic fluxes in continuous cultures of *Penicillium chrysogenum*. *J. Biotechnol.* **45**:149-164
- (51) L. H. Christensen; J. Marcher; U. Schulze; M. Carlsen; R. W. Min; J. Nielsen; J. Villadsen (1996)  
Semi-on-line analysis for fast and precise monitoring of bioreaction processes. *Biotechnol.* **52**:237-247
- (52) C. M. Henriksen; S. S. Holm; H. S. Jørgensen; J. Nielsen; J. Villadsen (1997)  
Kinetic studies of the carboxylation of 6-amino-penicillanic acid to 8-hydroxy-penicillic acid. *Proc. Biochem.* **32**:85-91
- (53) M. Carlsen; J. Nielsen; J. Villadsen (1996)  
Growth and  $\alpha$ -amylase production of *Aspergillus oryzae* during continuous cultivations. *J. Biotechnol.* **45**:81-93
- (54) Rong Wei Min; J. Nielsen; J. Villadsen (1996)  
On-line monitoring of glucose and penicillin by Sequential Injection Analysis. *Anal. Chim. Acta* **320**:199-205
- (55) P. de N. Pissarra; J. Nielsen; M. J. Bazin (1996)  
Pathway kinetics and metabolic control analysis of a high-yielding strain of *Penicillium chrysogenum* during fed-batch cultivations. *Biotechnol. Bioeng.* **51**:168-176
- (56) U. Schulze; G. Liden; J. Nielsen; J. Villadsen (1996)  
Physiological effects of nitrogen starvation in an anaerobic batch culture of *Saccharomyces cerevisiae*. *Microbiology* **142**:2299-2310
- (57) J. Nielsen; H. S. Jørgensen (1996)  
Kinetic model for the penicillin biosynthetic pathway in *Penicillium chrysogenum*. *Control Eng. Prac.* **4**:765-771
- (58) C. J. L. Klein, L. Olsson, B. Rønnow, J. D. Mikkelsen, J. Nielsen (1996)  
Alleviation of glucose repression on maltose metabolism by *MIG1* disruption in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **62**:4441-4449
- 1997**
- (59) J. Nielsen (1997)  
Metabolic control analysis of biochemical pathways based on a thermokinetic description of reaction rates. *Biochem. J.* **321**:133-138
- (60) T. L. Nissen; U. Schulze; J. Nielsen; J. Villadsen (1997)  
Flux distributions in anaerobic, glucose limited continuous cultures of *Saccharomyces cerevisiae*. *Microbiology* **143**:203-218
- (61) K. Schmidt; M. Carlsen; J. Nielsen; J. Villadsen (1997)  
Modelling isotopomer distributions in biochemical networks Using isotopomer mapping matrices. *Biotechnol. Bioeng.* **55**:831-840
- (62) T. C. Zangirolami; C. L. Johansen; J. Nielsen; S. B. Jørgensen (1997)  
Simulation of penicillin production in fed-batch cultivations using a morphologically structured model. *Biotechnol. Bioeng.* **56**:593-604
- (63) P. de N. Pissarra; J. Nielsen (1997)  
Thermodynamics of metabolic pathways for penicillin production: Analysis of thermodynamic feasibility and free energy changes during fed-batch cultivations. *Biotechnol. Prog.* **13**:156-165
- (64) M. Carlsen, K. V. Jøcumsen, C. Emborg; J. Nielsen (1997)  
Modeling the growth and proteinase A production in continuous cultures of recombinant *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **55**:447-454
- (65) P. Krabben; J. Nielsen; M. L. Michelsen (1997)  
Analysis of single hyphal growth and fragmentation in submerged cultures using a population model. *Chem. Eng. Sci.* **52**:2641-2652
- (66) A. B. Spøhr; M. Carlsen; J. Nielsen; J. Villadsen (1997)  
Morphological characterization of recombinant strains of *Aspergillus oryzae* producing  $\alpha$ -amylase during batch cultivations. *Biotechnol. Letters* **19**:257-261
- (67) L. Olsson; M. E. Larsen; B. Rønnow; J. D. Mikkelsen; J. Nielsen (1997)

**1998**

- (68) **J. Nielsen** (1998)  
Metabolic Engineering: Techniques for analysis of targets for genetic manipulations. *Biotechnol. Bioeng.* **58**:125-132
  - (69) A. Spohr; M. Carlsen; **J. Nielsen**; J. Villadsen (1998)  
 $\alpha$ -Amylase production in recombinant *Aspergillus oryzae* during fed-batch and continuous cultivations. *J. Ferment. Bioeng.* **86**:49-56
  - (70) H. B. Aa. Theilgaard; C. M. Henriksen; K. Kristiansen; **J. Nielsen** (1997)  
Purification and characterization of  $\delta$ -(L- $\alpha$ -aminoadipyl)-L-cysteinyl-D-valine synthetase (ACVS) from *Penicillium chrysogenum*. *Biochem. J.* **327**:185-191
  - (71) K. Schmidt; A. Marx; A. A. de Graaf; W. Wiechert; H. Sahm; **J. Nielsen**; J. Villadsen (1998)  
 $^{13}\text{C}$  Tracer experiments and metabolite balancing for metabolic flux analysis: Comparing two approaches. *Biotechnol. Bioeng.* **58**:254-257
  - (72) T. Agger; A. B. Spohr; M. Carlsen; **J. Nielsen** (1998)  
Growth and product formation of *Aspergillus oryzae* during submerged cultivations: Verification of a morphologically structured model using fluorescent probes. *Biotechnol. Bioeng.* **57**:321-329
  - (73) C. M. Henriksen; **J. Nielsen**; J. Villadsen (1997)  
Influence of dissolved oxygen concentration on the penicillin biosynthetic pathway in steady state cultures of *Penicillium chrysogenum*. *Biotechnol. Prog.* **13**:776-782
  - (74) A. B. Spohr; C. Dam Mikkelsen; M. Carlsen; **J. Nielsen**; J. Villadsen (1998)  
On-line study of fungal morphology during submerged growth in a small flow-through cell. *Biotechnol. Bioeng.* **58**:541-553
  - (75) K. Schmidt; **J. Nielsen**; J. Villadsen (1999)  
Quantitative analysis of metabolic fluxes in *E. coli*, using 2 dimensional NMR spectroscopy and complete isotopomer models. *J. Biotechnol.* **71**:175-190
  - (76) K. K. Klein; L. Olsson; B. Rennow; J. D. Mikkelsen; **J. Nielsen** (1998)  
Glucose and maltose metabolism in *MIG1*-disrupted and *MAL*-constitutive strains of *Saccharomyces cerevisiae*. *Food Technol. Biotechnol.* **35**:287-292
  - (77) H. P. Smits; A. Cohen; T. Buttler; **J. Nielsen**; L. Olsson (1998)  
Clean-up and analysis of sugar phosphates in biological extracts by using solid phase extraction and anion-exchange chromatography with pulsed amperometric detection. *Anal. Biochem.* **261**:36-42
  - (78) J. Holmalahti; O. Raatikainen; A. von Wright; H. Laatsch; A. Spohr; O. K. Lyngberg; **J. Nielsen** (1998)  
Production of dihydroabikoviromycin by *Streptomyces anulatus*. Production parameters and chemical characterization of genotoxicity. *J. Appl. Microbiol.* **85**:61-68
  - (79) C. J. L. Klein; L. Olsson; **J. Nielsen** (1998)  
Nitrogen-limited continuous cultivations as a tool to quantify glucose control in *Saccharomyces cerevisiae*. *Enz. Microbiol. Technol.* **23**:91-100
  - (80) C. M. Henriksen; **J. Nielsen**; J. Villadsen (1998)  
High exogenous concentrations of phenoxacetic acid are crucial for a high penicillin V productivity in *Penicillium chrysogenum*. *Microbiol.* **144**:2001-2006
  - (81) C. M. Henriksen; **J. Nielsen**; J. Villadsen (1998)  
Modelling of the protonophoric uncoupling by phenoxacetic acid of the plasma membrane of *Penicillium chrysogenum*. *Biotechnol. Bioeng.* **60**:761-767
  - (82) J. Dynesen; H. P. Smits; L. Olsson; **J. Nielsen** (1998)  
Carbon catabolite repression of invertase during batch cultivations of *Saccharomyces cerevisiae*: The role of glucose, fructose, and mannose. *Appl. Microbiol. Biotechnol.* **50**:579-582
  - (83) S. Ostergaard; H. B. Aa. Theilgaard; **J. Nielsen** (1998)  
Identification and purification of O-acetyl-L-serine sulphydrylase in *Penicillium chrysogenum*. *Appl. Microbiol. Biotechnol.* **50**:663-668
- 1999**
- (84) A. Aleksenko; W. Liu; Z. Gokovic; **J. Nielsen**; J. Piskur (1999)  
Structural and transcriptional analysis of the *pyrABCN*, *pyrC*, and *PyrF* genes in *Aspergillus nidulans* and the evolutionary origin of fungal dihydroorotases. *Mol. Microbiol.* **33**:599-611
  - (85) H. Pedersen; M. Carlsen; **J. Nielsen** (1999)  
Identification of enzymes and quantification of metabolic fluxes in the wild type and in a recombinant strain of *Aspergillus oryzae* strain. *Appl. Environ. Microbiol.* **65**:11-19
  - (86) A. L. Santerre Henriksen; S. Even; C. Müller; P. J. Punt; C. A. M. J. J. van den Hondel; **J. Nielsen** (1999)  
Study of the glucosylase promoter in *Aspergillus niger* using green fluorescence protein. *Microbiol.* **145**:729-734
  - (87) T. Christiansen; A. Spohr; **J. Nielsen** (1999)  
On-line study of growth kinetics of single hyphae of *Aspergillus oryzae* in a flow-through cell. *Biotechnol. Bioeng.* **63**:147-153
  - (88) A. Heydorn; T. Suhr-Jessen; **J. Nielsen** (1999)

- Growth and production kinetics of a teicoplanin producing strain of *Actinoplanes teicomyceticus*. *J. Antibio.* **52**:40-44
- (89) H. Aae, Theilgaard, **J. Nielsen** (1999)  
Metabolic control analysis of the penicillin biosynthetic pathway: The influence of the LLD-ACV:bisACV ratio on the flux control. *Antonie van Leeuwenhoek* **75**:145-154
- (90) C. J. L. Klein, J. J. Rasmussen, B. Rønnow, L. Olsson, **J. Nielsen** (1999)  
Investigation of the impact of *MIG1* and *MIG2* on the physiology of *Saccharomyces cerevisiae*. *J. Biotechnol.* **68**:197-212
- (91) M. Anderlund; T. L. Nissen; **J. Nielsen**; J. Villadsen; J. Rydstrom; B. Hahn-Hägerdal; M. C. Kiehlbrandt (1999)  
Expression of the *E. coli* *pntA* and *pntB* genes encoding nicotinamide nucleotide transhydrogenase in *Saccharomyces cerevisiae* and its effect on product formation during anaerobic glucose fermentation. *Appl. Environ. Microbiol.* **65**:2333-2340
- (92) K. Schmidt; L. C. Norregaard; B. Pedersen; A. Meissner; J. Ø. Duus; **J. Nielsen**; J. Villadsen (1999)  
Quantification of intracellular metabolic fluxes from fractional enrichment measurements and isotopomer analysis of <sup>13</sup>C labelled biomass components. *Metabol. Eng.* **1**:166-179
- (93) B. Christensen; **J. Nielsen** (1999)  
Isotopomer analysis using GC-MS. *Metabol. Eng.* **1**:282-290
- (94) A. Santerre Henriksen, M. Carlsen; H. de Bang, **J. Nielsen** (1999)  
Kinetics of  $\alpha$ -amylase secretion in *Aspergillus oryzae*. *Biotechnol. Bioeng.* **65**:76-82
- (95) B. Rønnow; L. Olsson; **J. Nielsen**; J. D. Mikkelsen (1999)  
Deroppression of galactose metabolism in melibiase producing baker's and distillers' yeast *J. Biotechnol.* **72**:213-228
- 2000**
- (96) H. Pedersen; **J. Nielsen** (2000)  
The influence of nitrogen sources on the  $\alpha$ -amylase productivity of *Aspergillus oryzae* in continuous Culture. *Appl. Microb. Biotechnol.* **53**:278-281
- (97) T. Agger; **J. Nielsen** (1999)  
Genetically structured modelling of protein production in filamentous fungi. *Biotechnol. Bioeng.* **66**:164-170
- (98) C. Müller, A. Spohr; **J. Nielsen** (2000)  
Role of substrate concentration on mitosis and hyphal extension of *Aspergillus*. *Biotechnol. Bioeng.* **67**:390-397
- (99) H. Pedersen; B. Christensen; C. Hjort; **J. Nielsen** (2000)  
Construction and characterization of an oxalic acid non-producing strain of *Aspergillus niger*. *Metabol. Eng.* **2**:34-41
- (100) H. Pedersen; M. Beyer; **J. Nielsen** (2000)  
Glucosylase production in batch, chemostat and fed-batch cultivations by an industrial strain of *Aspergillus niger*. *Appl. Microbiol. Biotechnol.* **53**:272-277
- (101) S. Ostergaard; C. Roca; B. Rønnow; **J. Nielsen**; L. Olsson (2000)  
Physiological studies in aerobic batch cultivations of *Saccharomyces cerevisiae* strains harbouring the *MEL1* gene. *Biotechnol. Bioeng.* **68**:252-259
- (102) H. Pedersen; C. Hjort; **J. Nielsen** (2000)  
Cloning and characterization of *oah*, the gene encoding oxaloacetate hydrolase from *Aspergillus niger*. *Mol. Gen. Genetics.* **263**:281-286
- (103) L. Olsson; **J. Nielsen** (2000)  
The role of metabolic engineering in the improvement of *Saccharomyces cerevisiae*: Utilization of industrial media. *Enz. Microbiol. Technol.* **26**:785-792
- (104) T. L. Nissen; C. W. Hamann; M. C. Kiehlbrandt; **J. Nielsen**; J. Villadsen (2000)  
Anaerobic and batch cultivations of *Saccharomyces cerevisiae* mutants impaired in glycerol synthesis. *Yeast* **16**:463-474
- (105) T. L. Nissen; M. C. Kiehlbrandt; **J. Nielsen**; J. Villadsen (2000)  
Optimisation of ethanol production in *Saccharomyces cerevisiae* by metabolic engineering of the ammonia assimilation. *Metabol. Eng.* **2**:69-77
- (106) A. Heydorn; B. O. Pedersen; J. Ø. Duus, S. Bergmann; T. Suhr-Jessen; **J. Nielsen** (2000)  
Biosynthetic studies of the glycopeptide teicoplanin by <sup>1</sup>H and <sup>13</sup>C NMR. *J. Biol. Chem.* **275**:6201-6206
- (107) J. P. van Dijken *et al.* (2000)  
An inter-laboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. *Enz. Microb. Technol.* **26**:706-714
- (108) T. C. Zangirolami; M. Carlsen; **J. Nielsen**; S. B. Jørgensen (2000)  
Selection and characterization of a high  $\alpha$ -amylase-producing variant in glucose-limited continuous cultures of *Aspergillus oryzae*. *Myc. Res.* **104**:1241-1249
- (109) B. Christensen; **J. Nielsen** (2000)  
Metabolic network analysis on *Penicillium chrysogenum* using <sup>13</sup>C-labelled glucose. *Biotechnol. Bioeng.* **68**:652-659
- (110) B. Christensen; J. Thykær; **J. Nielsen** (2000)  
Metabolic characterization of high and low yielding strains of *Penicillium chrysogenum*. *Appl. Microbiol. Biotechnol.* **54**:212-217
- (111) **J. Nielsen** (2000)

- The role of metabolic engineering in the improvement of industrial processes. *Thai J. Biotechnol.* **2**:14-25
- (112) H. P. Smits; J. Hauf; S. Müller; T. J. Hobley; F. K. Zimmermann; B. Hahn-Hägerdal; J. Nielsen; L. Olsson (2000) Simultaneous over-expression of enzymes of the lower part of glycolysis can enhance the fermentative capacity of *Saccharomyces cerevisiae*. *Yeast* **16**:1325-1334
  - (113) H. Aae Theilgaard; M. van den Berg; C. Mulder; R. A. L. Bovenberg; J. Nielsen (2000) Quantitative analysis of *Penicillium chrysogenum* Wis54-1255 transformants over-expressed in the penicillin biosynthetic genes. *Biotechnol. Bioeng.* **72**:379-388
  - (114) L. F. Bautista; M. Hentzer; A. Santerre Henriksen; A. Aleksenko; J. Nielsen (2000) Antisense silencing of the *creA* gene in *Aspergillus nidulans*. *Appl. Environ. Microbiol.* **66**:4579-4581
- 2001**
- (115) T. Agger; A. Spohr; J. Nielsen (2001)  $\alpha$ -Amylase production in high cell density submerged cultivations of *Aspergillus oryzae* and *Aspergillus nidulans*. *Appl. Microbiol. Biotechnol.* **55**:81-84
  - (116) S. Ostergaard; L. Olsson; M. Johnston; J. Nielsen (2000) Increasing galactose consumption by *Saccharomyces cerevisiae* through metabolic engineering of the *GAL* gene regulatory network. *Nature Biotechnol.* **18**:1283-1286
  - (117) E. Jonsbu; T. Elingsen; J. Nielsen (2001) Effects of nitrogen sources on cell growth and production of nystatin by *Streptomyces noursei* J. *Antibiot.* **53**:1354-1362
  - (118) M. McIntyre; J. Dynesen; J. Nielsen (2001) Morphological characterization of *Aspergillus nidulans*: Growth, septation and fragmentation. *Microbiol.* **147**:239-246
  - (119) A. K. Gombert; M. M. dos Santos; B. Christensen; J. Nielsen (2001) Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* at different conditions of glucose repression. *J. Bacteriol.* **183**:1441-1451
  - (120) S. Ostergaard; L. Olsson; J. Nielsen (2001) *In vivo* dynamics of galactose metabolism in *Saccharomyces cerevisiae*: Metabolic fluxes and metabolite levels. *Biotechnol. Bioeng.* **73**:412-425
  - (121) S. Ostergaard; K. O. Walløe; C. S. G. Gomes; L. Olsson; J. Nielsen (2001) The impact of *GAL6*, *GAL80*, and *MIG1* on glucose control of the *GAL* system in *Saccharomyces cerevisiae*. *FEMS Yeast Research.* **1**:47-55
  - (122) T. L. Nissen; M. Anderlund; M. C. Kiehlbrandt; J. Nielsen; J. Villadsen (2001) Expression of a cytoplasmic transhydrogenase in *Saccharomyces cerevisiae* results in formation of 2-oxoglutarate due to depletion of the NADPH pool. *Yeast* **18**:19-32
  - (123) K. Møller; L. D. Tidemand; J. R. Winther; L. Olsson; J. Piskur; J. Nielsen (2001) Production of a heterologous proteinase A by *Saccharomyces kluyveri*. *Appl. Microbiol. Biotechnol.* **57**:216-219
  - (124) B. Christensen; T. Christiansen; A. K. Gombert; J. Thykær; J. Nielsen (2001) A simple and robust method for estimation of the flux through the pentose phosphate pathway in micro-organisms. *Biotechnol. Bioeng.* **74**:517-523
  - (125) E. Jonsbu; B. Christensen; J. Nielsen (2001) Changes in *in vivo* fluxes through central metabolic pathways during production of nystatin by *Streptomyces noursei* in batch culture. *Appl. Microbiol. Biotechnol.* **56**:93-100
  - (126) M. Carlsen; J. Nielsen (2001) Influence of carbon source on  $\alpha$ -amylase production by *Aspergillus oryzae*. *Appl. Microbiol. Biotechnol.* **57**:346-349
  - (127) T. Agger; J. B. Petersen; S. M. O'Connor; R. L. Murphy; J. M. Kelly; J. Nielsen (2002) Physiological characterisation of recombinant *Aspergillus nidulans* strains with different *creA* genotypes expressing *A. oryzae*  $\alpha$ -amylase. *J. Biotechnol.* **92**:279-285
  - (128) T. C. Zangirolami; M. Carlsen; J. Nielsen; S. B. Jørgensen (2002) Growth and enzyme production during continuous cultures of a high amylase-producing variant of *Aspergillus oryzae*. *Brazilian J. Chem. Eng.* **19**:55-68
  - (129) J. Robin; M. Jacobsen; M. Beyer; H. Noorman; J. Nielsen (2001) Physiological characterisation of *Penicillium chrysogenum* strains expressing the expandase gene from *Streptomyces clavuligerus* during batch cultivations. Growth and adipoyl-7-ADCA production. *Appl. Microbiol. Biotechnol.* **57**:357-362
  - (130) K. Møller; B. Christensen; J. Förster; J. Piskur; J. Nielsen; L. Olsson (2001) Aerobic glucose metabolism of *Saccharomyces kluyveri*: Growth, metabolite production and quantification of metabolic fluxes. *Biotechnol. Bioeng.* **77**:186-193
- 2002**
- (131) T. Christiansen; B. Christensen; J. Nielsen (2002) Metabolic network analysis of *Bacillus clausii* on minimal and semi-rich medium using  $^{13}\text{C}$ -labeled glucose. *Metabolic Eng.* **4**:159-169
  - (132) J. Thykær; B. Christensen; J. Nielsen (2002) Metabolic network analysis of an adipoyl-7-ADCA producing strain of *Penicillium chrysogenum*: Elucidation of adipate degradation. *Metabolic Eng.* **4**:151-158



- (133) M. McIntyre; J. Breum; J. Arnau; J. Nielsen (2002)  
Growth physiology and dimorphism of *Mucor circinelloides* (syn. *racemosus*) during submerged batch cultivation. *Appl. Microbiol. Biotechnol.* **58**:495-502
  - (134) T. Christiansen; J. Nielsen (2002)  
Growth energetics of an alkaline serine protease producing strain of *Bacillus clausii* during continuous cultivation. *Bioproc. Biosystems Eng.* **24**:329-339
  - (135) B. Christensen; J. Nielsen (2002)  
Reciprocal  $^{13}\text{C}$ -labeling: A method for investigating the catabolism of co-substrates. *Biotechnol. Prog.* **18**:163-166
  - (136) C. Müller; M. McIntyre; K. Hansen; J. Nielsen (2002)  
Metabolic engineering of the morphology of *Aspergillus oryzae* by altering chitin synthesis. *Appl. Environ. Microbiol.* **68**:1827-1836
  - (137) A. G. Vara; A. Hochkoeppler; J. Nielsen; J. Villadsen (2002)  
Production of teicoplanin by *Actinoplanes teichomyceticus* in continuous fermentation. *Biotechnol. Bioeng.* **77**:589-598
  - (138) J. Nielsen; L. Olsson (2002)  
An expanded role for microbial physiology in metabolic engineering and functional genomics: Moving towards systems biology. *FEMS Yeast Res.* **2**:175-181
  - (139) N. B. S. Jensen; B. Christensen; J. Nielsen; J. Villadsen (2002)  
The simultaneous biosynthesis and uptake of amino acids by *Lactococcus lactis* studied by  $^{13}\text{C}$  labelling experiments. *Biotechnol. Bioeng.* **78**:11-16
  - (140) K. Möller; C. Bro; J. Piskur; J. Nielsen; L. Olsson (2002)  
Steady-state and transient-state analysis of aerobic fermentation in *Saccharomyces kluyveri*. *FEMS Yeast Res.* **2**:233-244
  - (141) T. Christiansen; J. Nielsen (2002)  
Production of extracellular protease and glucose uptake in *Bacillus clausii* in steady state and transient continuous cultures. *J. Biotechnol.* **97**:265-273
  - (142) J. Förster; A. K. Gombert; J. Nielsen (2002)  
A functional genomics approach using metabolomics and *in silico* pathway analysis. *Biotechnol. Bioeng.* **79**:703-712
  - (143) J. Zaldivar; A. Borges; B. Johansson; H. P. Smits; S. G. Villas-Boas; J. Nielsen; L. Olsson (2002)  
Fermentation performance and intracellular metabolite patterns in laboratory and industrial xylose fermenting *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **59**:436-442
  - (144) C. Müller; C. M. Hjort; K. Hansen; J. Nielsen (2002)  
Altering the expression of two chitin synthase genes differentially affects the growth and morphology of *Aspergillus oryzae*. *Microbiology* **148**:4025-4033
  - (145) B. Christensen; A. K. Gombert; J. Nielsen (2002)  
Analysis of flux estimates based on  $^{13}\text{C}$ -labeling experiments. *Eur. J. Biochem.* **269**:2795-2800
  - (146) E. Jonsbu; M. McIntyre; J. Nielsen (2002)  
The influence of carbon sources and morphology on nystatin production by *Streptomyces noursei*. *J. Biotechnol.* **95**:133-144
  - (147) M. D. W. Piper; P. Daran-Lapujade; C. Bro; B. Regenber; S. Knudsen; J. Nielsen; J. T. Pronk (2002)  
Reproducibility of oligonucleotide microarray transcriptome analysis: an interlaboratory comparison using chemostat cultures of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **277**:37001-37008
- 2003**
- (148) C. Müller; K. Hansen; P. Szabo; J. Nielsen (2003)  
Effect of chitin synthesis on mycelial morphology and culture viscosity in *Aspergillus oryzae*. *Biotechnol. Bioeng.* **81**:525-534
  - (149) T. Christiansen; S. Michaelsen; M. Wümpelmann; J. Nielsen (2003)  
Production of Savinase and population viability of *Bacillus clausii* during high cell density fed-batch cultivations. *Biotechnol. Bioeng.* **83**:344-352
  - (150) J. Robin; G. Lettier; M. McIntyre; H. Noorman; J. Nielsen (2003)  
Continuous cultivations of *Penicillium chrysogenum* strain expressing the expandase gene from *Streptomyces clavuligerus*: Growth yields and morphological characterisation. *Biotechnol. Bioeng.* **83**:353-360
  - (151) J. Robin; P. Bruheim; M. L. Nielsen; H. Noorman; J. Nielsen (2003)  
Continuous cultivations of *Penicillium chrysogenum* strain expressing the expandase gene from *Streptomyces clavuligerus*: Kinetics of adipoyl-7-aminodeacetoxycephalosporanic acid and by-product formation. *Biotechnol. Bioeng.* **83**:361-368
  - (152) J. Förster; I. Famili; P. Fu; B. Ø. Palsson; J. Nielsen (2003)  
Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res.* **13**:244-253
  - (153) T. L. Lübbehüsen; J. Nielsen; M. McIntyre (2003)  
Characterization of the *Mucor circinelloides* life cycle by on-line image analysis. *J. Appl. Microbiol.* **95**:1152-1160
  - (154) J. Robin; S. Bonneau; D. Schipper; H. Noorman; J. Nielsen (2003)  
Influence of the adipate and dissolved oxygen concentrations on the  $\beta$ -lactams production during continuous cultivations of a *Penicillium chrysogenum* strain expressing the expandase gene from *Streptomyces clavuligerus*. *Metabol. Eng.* **5**:42-48

- (155) N. Gunnarsson; P. Bruheim; J. Nielsen (2003)  
Production of the glycopeptide antibiotic A40926 by *Nonomuria* sp. ATCC 39727: influence of medium composition in batch fermentation. *J. Ind. Microbiol. Biotechnol.* **30**:150-156
  - (156) J. Dynesen; J. Nielsen (2003)  
Surface hydrophobicity of *Aspergillus nidulans* conidiophores and its role in pellet formation. *Biotechnol. Prog.* **19**:1049-1052
  - (157) M. M. dos Santos; A. K. Gombert; B. Christensen; L. Olsson; J. Nielsen (2003)  
Identification of *in vivo* enzyme activities in the co-metabolism of glucose and acetate by *Saccharomyces cerevisiae* using  $^{13}\text{C}$ -labeled substrates. *Eukar. Cell.* **2**:599-608
  - (158) J. Förster; I. Famili; B. Ø. Palsson; J. Nielsen (2003)  
Large-scale Evaluation of *in silico* Gene Deletions in *Saccharomyces cerevisiae*. *Omics J. Integrative Biol.* **7**:193-202
  - (159) T. L. Lübbehüsen; J. Nielsen; M. McIntyre (2003)  
Morphology and physiology of the dimorphic fungus *Mucor circinelloides* (syn. *M. racemosus*) during anaerobic growth. *Mycol. Res.* **107**:223-230
  - (160) C. Roca; J. Nielsen; L. Olsson (2003)  
Metabolic engineering of ammonium assimilation in xylose fermenting *Saccharomyces cerevisiae* improves ethanol production. *Appl. Env. Microbiol.* **69**:4732-4736
  - (161) W. Prathumpai; M. McIntyre; J. Nielsen (2004)  
The effect of CreA in glucose and xylose catabolism in *Aspergillus nidulans*. *Appl. Microbiol. Biotechnol.* **63**:748-753
  - (162) C. Bro; B. Regenber; G. Lagniel; J. Labarre; M. Montero-Lomeli; J. Nielsen (2003)  
Transcriptional, proteomics and metabolic response to lithium in galactose-grown yeast cells. *J. Biol. Chem.* **278**:32141-32149
  - (163) W. Prathumpai; J. B. Gabelgaard; P. Wanchanthuek; P. J. I. van de Vondervoort; M. J. L. de Groot; M. McIntyre; J. Nielsen (2003)  
Metabolic control analysis of xylose catabolism in *Aspergillus*. *Biotechnol. Prog.* **19**:1136-1141
  - (164) H. David; M. Åkesson; J. Nielsen (2003)  
Reconstruction of the central carbon metabolism of *Aspergillus niger*. *Eur. J. Biochem.* **270**: 4243-4253
  - (165) M. M. dos Santos; G. Thygesen; P. Kötter; L. Olsson; J. Nielsen (2003)  
Aerobic physiology of redox-engineered *Saccharomyces cerevisiae* strains modified in the ammonium assimilation for increased NADPH availability. *FEMS Yeast Res.* **4**:59-68
  - (166) I. Famili; J. Förster; J. Nielsen; B. Palsson (2003)  
*Saccharomyces cerevisiae* phenotypes can be predicted using constraint based analysis of a genome-scale reconstructed metabolic network. *Proc. Nat. Acad. Sci.* **100**:13134-13139
  - (167) J. Dynesen; J. Nielsen (2003)  
Branching is coordinated with mitosis in growing hyphae of *Aspergillus nidulans*. *Fungal Gen. Biol.* **40**:15-24
  - (168) S. G. Villas-Boas; D. G. Delicado; M. Åkesson; J. Nielsen (2003)  
Simultaneous analysis of amino and nonamino organic acids as methyl chloroformate derivatives using gas chromatography-mass spectrometry. *Anal. Biochem.* **322**:134-138
  - (169) J. Nielsen (2003)  
It is all about metabolic fluxes. *J. Bacteriol.*, **185**:7031-7035
- 2004**
- (170) T. L. Lübbehüsen, J. Nielsen, M. McIntyre (2004)  
Aerobic and anaerobic ethanol production by *Mucor circinelloides* during submerged growth. *Appl. Microbiol. Biotechnol.* **63**:543-548
  - (171) T. L. Lübbehüsen, V. G. Polo, S. Rossi, J. Nielsen, S. Moreno, M. McIntyre, J. Arnau (2004)  
Protein kinase A is involved in the control of morphology and branching during aerobic growth in *Mucor circinelloides*. *Microbiol.* **150**:143-150
  - (172) C. Bro; B. Regenber; J. Nielsen (2004)  
Genome-wide transcriptional response of a *Saccharomyces cerevisiae* strain with an altered redox metabolism. *Biotechnol. Bioeng.* **85**:269-276
  - (173) W. Prathumpai; M. McIntyre; J. Nielsen (2004)  
The effect of CreA in glucose and xylose catabolism in *Aspergillus nidulans*. *Appl. Microbiol. Biotechnol.* **63**:748-753
  - (174) C. Bro; J. Nielsen (2004)  
Impact of 'Omics' analysis on inverse metabolic engineering. *Met. Eng.* **6**:204-211
  - (175) K. Möller; C. Bro; R. B. Langkjaer; J. Piskur; J. Nielsen; L. Olsson (2004)  
Pyruvate decarboxylases from the petite-negative yeast *Saccharomyces kluyveri*. *Mol. Genet. Genomics* **270**:558-568
  - (176) T. Grotkjær; M. Åkesson; B. Christensen; A. K. Gombert; J. Nielsen (2004)  
Impact of transamination reactions and protein turnover on labeling dynamics in  $^{13}\text{C}$ -labelling experiments. *Biotechnol. Bioeng.* **86**:209-216
  - (177) M. Åkesson; J. Förster; J. Nielsen (2004)  
Integration of gene expression data into genome-scale metabolic models. *Met. Eng.* **6**:285-293

- (178) G. G. Larsen; K. F. Appel, A.-M. Wolff, J. Nielsen; J. Arnau (2004) Characterization of the *Mucor circinelloides* regulated promoter *gpd1P* and construction of a constitutive derivative. *Curr. Genet.* **45**:225-234
  - (179) N. Gunnarsson; U. H. Mortensen; M. Sosio; J. Nielsen (2004) Identification of the Entner-Doudoroff pathway in an antibiotic producing actinomycete species. *Mol Microbiol* **52**: 895-902
  - (180) N. Eckert-Boulet; P. S. Nielsen; C. Friis; M. M. dos Santos; J. Nielsen; M. C. Kiehlbrandt; B. Regenberg (2004) Transcriptional profiling of extracellular amino acid sensing in *Saccharomyces cerevisiae* and the role of *Stp1p* and *Stp2p*. *Yeast* **21**: 635-648
  - (181) S. L. Westergaard; C. Bro; L. Olsson; J. Nielsen (2004) Elucidation of the role of *Grr1p* on glucose sensing by *Saccharomyces cerevisiae* through genome-wide transcription analysis. *FEMS Yeast Res.* **5**:193-204
  - (182) V. Raghevedran; A. K. Gombert; B. Christensen; P. Kötter; J. Nielsen (2004) Phenotypic characterization of glucose repression mutants of *Saccharomyces cerevisiae* using experiments with  $^{13}\text{C}$ -labelled glucose. *Yeast* **21**:769-779
  - (183) T. Grotkjær; J. Nielsen (2004) Enhancing yeast transcription analysis through integration of heterogeneous data. *Cur. Genomics* **5**:673-686
  - (184) M. M. dos Santos; V. Raghevedran; P. Kötter; L. Olsson; J. Nielsen (2004) Manipulation of the malic enzyme in *Saccharomyces cerevisiae* for increasing NADPH production capacity aerobically in different cellular compartments. *Met. Eng.* **6**:352-363
  - (185) B. Regenberg; U. Krühne; M. Beyer; L. H. Pedersen; M. Simón; O. R.T. Thomas; J. Nielsen; T. Ahl (2004) Use of laminar flow patterning for miniaturised biochemical assays. *Lab. Chip* **4**:654-657
  - (186) W. Prathumpai; S. J. Flitter; M. McIntyre; J. Nielsen (2004) Lipase production by recombinant strains of *Aspergillus niger* expressing a lipase-encoding gene from *Thermomyces lanuginosus*. *Appl. Microbiol. Biotechnol.* **65**:714-719
  - (187) C. Cannizzaro; B. Christensen; J. Nielsen; U. von Stockar (2004) Metabolic network analysis on *Phaffia rhodozyma* yeast using  $^{13}\text{C}$ -labeled glucose and gas chromatography-mass spectrometry. *Met. Eng.* **6**:340-351
  - (188) N. Gunnarsson; P. Bruheim; J. Nielsen (2004) Glucose metabolism in the antibiotic producing Actinomycete *Nonomuraea* sp. *Biotechnol. Bioeng.* **88**: 652-663
- 2005**
- (189) G. Panagiotou, S. G. Villas-Boas, P. Christakopoulos, J. Nielsen, L. Olsson (2005) Intracellular metabolite profiling of *Fusarium oxysporum* converting glucose to ethanol. *J. Biotechnol.* **115**:425-434
  - (190) J. Smedsgaard, J. Nielsen (2005) Metabolite profiling of fungi and yeast: from phenotype to metabolome by MS and informatics. *J. Exp. Bot.* **56**:273-286
  - (191) K. R. Patil, J. Nielsen (2005) Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proc. Nat. Acad. Sci.* **102**:2685-2689
  - (192) I. Borodina, C. Schöller, A. Eliasson, J. Nielsen (2005) Metabolic network analysis of *Streptomyces tenebrarius*, a *Streptomyces* with Entner-Doudoroff pathway. *Appl. Environ. Microbiol.* **71**:2294-2302
  - (193) S. G. Villas-Bôas; M. Åkesson; J. Nielsen (2005) Biosynthesis of glyoxylate from glycine in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **5**:703-709
  - (194) S. G. Villas-Bôas; J. F. Moxley; M. Åkesson; G. Stephanopoulos; J. Nielsen (2005) High-throughput metabolic state analysis: The missing link in integrated functional genomics of yeasts. *Biochem. J.* **388**:669-677
  - (195) N. Eckert-Boulet; B. Regenberg; J. Nielsen (2005) *Grr1p* is required for transcriptional induction of amino acid permease genes and proper transcriptional regulation of genes in carbon metabolism of *Saccharomyces cerevisiae*. *Curr. Genet.* **47**:139-149
  - (196) T. Seker; K. Möller; J. Nielsen (2005) Analysis of acyl CoA ester intermediates of the mevalonate pathway in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **67**:199-124
  - (197) I. Borodina; P. Krabben; J. Nielsen (2005) Genome-scale analysis of *Streptomyces coelicolor* A3(2) metabolism. *Genome Res.* **15**:820-829
  - (198) A. P. Oliveira; J. Nielsen; J. Förster (2005) Modeling *Lactococcus lactis* using a genome-scale flux model. *BMC Microbiol.* **5**:39
  - (199) C. Bro; S. Knudsen; B. Regenberg; L. Olsson; J. Nielsen (2005) Improvement of galactose uptake in *Saccharomyces cerevisiae* through overexpression of phosphoglucosyltransferase: Example of transcript analysis as a tool in inverse metabolic engineering. *Appl. Environ. Microbiol.* **71**:6465-6472
  - (200) H. David; A. M. Krogh; G. Roca; M. Åkesson; J. Nielsen (2005) CreA influences the metabolic fluxes of *Aspergillus nidulans* during growth on glucose and xylose. *Microbiol.* **151**:2209-2221

- (201) V. Raghevedran, J. Nielsen, L. Olsson (2005)  
Teaching microbial physiology using glucose repression phenomenon in Baker's yeast as an example. *Biochem. Mol. Biol. Edu.* **33**:404-410
- (202) M. C. Jewett, A. P. Oliveira; K. R. Patil; J. Nielsen (2005)  
The role of high-throughput transcriptome analysis in metabolic engineering. *Biotechnol. Bioproc. Eng.* **10**:385-399
- (203) T. Grotkjær, P. Christakopoulos; J. Nielsen ; L. Olsson (2005)  
Comparative metabolic network analysis of two xylose fermenting recombinant *Saccharomyces cerevisiae* strains. *Met. Eng.* **7**:437-444
- (204) K. R. Patil; I. Rocha; J. Förster; J. Nielsen (2005)  
Evolutionary programming as a platform for *in silico* metabolic engineering. *BMC Bioinformatics* **6**:308
- (205) S. G. Villas-Bôas; J. H. Pedersen; M. Åkesson; I. Smedsgaard; J. Nielsen (2005)  
Global metabolite analysis of yeast: evaluation of sample preparation methods. *Yeast* **22**:1155-1169
- 2006**
- (206) T. Grotkjær, O. Winther, B. Regenberg, J. Nielsen, L. K. Hansen (2006)  
Robust multi-scale clustering of large DNA microarray datasets with the consensus algorithm. *Bioinformatics* **22**:58-67
- (207) N. Eckert-Boulet; K. Larsson; B. Wu; P. Poulsen; B. Regenberg, J. Nielsen; M. Kielland-Brandt (2006)  
Deletion of *RTS1*, encoding a regulatory subunit of protein phosphatase 2A, results in constitutive amino acid signaling via increased Sit1p processing. *Eukaryotic Cell* **5**:174-179
- (208) C. Bro; B. Regenberg, J. Förster, J. Nielsen (2006)  
*In silico* aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Met. Eng.* **8**:102-111
- (209) R. Mustachi; S. Hohmann; J. Nielsen (2006)  
Yeast systems biology to unravel the network of life. *Yeast* **23**:227-238
- (210) J. Mogensen; H. B. Nielsen; G. Hofmann; J. Nielsen (2006)  
Transcription analysis using high-density micro-arrays of *Aspergillus nidulans* wild-type and *ΔcreA* mutant during growth on glucose and ethanol. *Fung. Gen. Bio.* **43**:593-603
- (211) V. Raghevedran; K. R. Patil; L. Olsson; J. Nielsen (2006)  
Hap4 is not essential for activation of respiration at low specific growth rates in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **281**:12308-12314
- (212) R. Usaité; K. R. Patil; T. Grotkjær; J. Nielsen; B. Regenberg (2006)  
Global transcriptional and physiological responses of *Saccharomyces cerevisiae* to ammonium, L-alanine or L-glutamine limitation. *Appl. Environ. Microbiol.* **72**:6194-6203
- (213) A. Diano; S. Bekker-Jensen, J. Dynesen, J. Nielsen (2006)  
Polyol synthesis in *Aspergillus niger*: Influence of oxygen availability, carbon and nitrogen sources on the metabolism. *Biotechnol. Bioeng.* **94**:899-908
- (214) B. Regenberg; T. Grotkjær, O. Winther, A. Fausbøll; M. Åkesson; C. Bro; L. K. Hansen, S. Brunak; J. Nielsen (2006)  
Growth-rate regulated genes have profound impact on interpretation of transcriptome profiling in *Saccharomyces cerevisiae*. *Genome Biol.* **7**:R107
- (215) H. David; G. Hofmann; A. P. Oliveira; H. Jarmer; J. Nielsen (2006)  
Metabolic network driven analysis of genome-wide transcription data from *Aspergillus nidulans*. *Genome Biol.* **7**:R108
- (216) T. Cakir; K. R. Patil; Z. I. Önsan; K. Ö. Ülgen; B. Kirdar; J. Nielsen (2006)  
Integration of metabolome data with metabolic networks reveals reporter reactions. *Mol. Systems Biol.* **2**:50
- 2007**
- (217) S. L. Westergaard; A. P. Oliveira; C. Bro; L. Olsson; J. Nielsen (2007)  
A systems biology approach to study glucose repression in the yeast *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **96**:134-145
- (218) G. N. Vemuri; M. A. Eiteman; J. E. McElwen; L. Olsson; J. Nielsen (2007)  
Increasing NADH oxidation reduces overflow metabolism in *Saccharomyces cerevisiae*. *Proc. Nat. Acad. Sci.* **104**:2402-2407
- (219) H. J. Pel et al. (2007)  
Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS513.88. *Nature Biotech.* **25**:221-231
- (220) S. Mass; S. G. Villas-Boas; M. E. Hansen; M. Åkesson; J. Nielsen (2007)  
A comparison of direct infusion MS and GC-MS for metabolic footprinting of yeast mutants. *Biotechnol. Bioeng.* **96**:1014-1022
- (221) S. Wattanachaisaereekul; A. Eliasson Lantz; M. L. Nielsen; O. S. Andresson; J. Nielsen (2007)  
Optimization of heterologous production of the polyketide 6-MSA in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **97**:893-900
- (222) C. J. Baxter; H. Redestig; N. Schauer; D. Repsilber; K. R. Patil; J. Nielsen; J. Selbig; J. Liu; A. R. Fernie; L. J. Sweetlove (2007)  
The metabolic response of heterotrophic *Arabidopsis* cells to oxidative stress. *Plant Phys.* **143**:312-325
- (223) K. Y. Arga; Z. I. Önsan; B. Kirdar; K. Ö. Ülgen; J. Nielsen (2007)

- Understanding signaling in yeast: Insights from network analysis. *Biotechnol. Bioeng.* 97:1246-1258
- (224) T. Cakir; B. Kirdar; Z. I. Onsan; K. O. Ulgen; J. Nielsen (2007)  
Effect of carbon source perturbations on transcriptional regulation of metabolic fluxes in *Saccharomyces cerevisiae*. *BMC Systems Biol.* 1:18
- (225) J. Thykaer; J. Nielsen (2007)  
Evidence, through C13-labelling analysis, of phosphoketolase activity in fungi. *Proc. Biochem.* 42:1050-1055
- (226) S. Meier; G. Panagiotou; L. Olsson; J. Nielsen (2007)  
Physiological characterization of xylose metabolism in *Aspergillus niger* under oxygen limited conditions. *Biotechnol. Bioeng.* 98:462-475
- (227) I. Nookaew; A. Meechai; C. Thammarongtham; K. Loateng; V. Ruanglek; S. Cheevadhanarak; J. Nielsen; S. Bhumirata (2007)  
Identification of flux regulation coefficients from elementary flux modes: A systems biology tool for analysis of metabolic networks. *Biotechnol. Bioeng.* 97:1535-1549
- (228) M. L. Nielsen; W. A. de Jongh; S. L. Meijer; J. Nielsen; U. Mortensen (2007)  
Transient marker system for iterative gene targeting of a prototrophic fungus. *Appl. Environ. Microbiol.* 73:7240-7245
- (229) M. J. van der Werf; R. Takors; J. Smedsgaard; J. Nielsen; T. Ferenci; J. C. Portais; C. Wittmann; M. Hooks; A. Tomassini; M. Oldiges; J. Postel; U. Sauer (2007)  
Standard reporting requirements for biological samples in metabolomics experiments: Microbial and in vitro biology experiments. *Metabolomics* 3:189-194
- 2008**
- (230) R. Usaite; J. Nielsen; L. Olsson (2008)  
Physiological characterization of glucose repression in the strains with *SNF1* and *SNF4* genes deleted. *J. Biotechnol.* 133:73-81
- (231) R. Usaite; J. Wohlschlegel; J. D. Venable; S. K. Park; J. Nielsen; L. Olsson; John R. Yates III (2008)  
Characterization of global yeast quantitative proteome data generated from the wild type and glucose repression *Saccharomyces cerevisiae* strains: the comparison of two quantitative algorithms. *J. Proteom Res.* 7:266-275
- (232) W. de Jongh; J. Nielsen (2007)  
Enhanced citrate production through gene insertion in *Aspergillus niger*. *Met. Eng.* 10:87-96
- (233) M. Asadollahi; J. Maury; K. Møller; K. F. Nielsen; M. Schalk; A. Clark; J. Nielsen (2008)  
Production of plant sesquiterpenes in *Saccharomyces cerevisiae*: Effect of *ERG9* repression on sesquiterpene biosynthesis. *Biotechnol. Bioeng.* 99:666-677
- (234) M. R. Andersen; M. L. Nielsen; J. Nielsen (2008)  
Metabolic model integration of the bibliome, genome, metabolome and reactome of *Aspergillus niger*. *Mol. Systems Biol.* 4:178
- (235) A. P. Oliveira; K. Patil; J. Nielsen (2008)  
Architecture of transcriptional regulatory circuits is knitted over the topology of bio-molecular interaction networks. *BMC Systems Biol.* 2:17
- (236) J. Thykaer; K. Rueksomtawin; H. Noorman; J. Nielsen (2008) NADPH-Dependent glutamate dehydrogenase in *Penicillium chrysogenum* is involved in regulation of  $\beta$ -lactam production. *Microbiol.* 154:1242-1250
- (237) H. David; I. Ozelik; G. Hofmann; J. Nielsen (2008) Analysis of *Aspergillus nidulans* metabolism at the genome-scale. *BMC Genomics* 9:163
- (238) W. A. de Jongh; C. Bro; S. Ostergaard; B. Regenber; L. Olsson; J. Nielsen (2008)  
The roles of galactitol, galactose-1-phosphate and phosphoglucomutase in galactose-induced toxicity in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 101:317-326
- (240) M. R. Andersen; W. Vongsangnak; G. Panagiotou; M. P. Salazar; L. Lehmann; J. Nielsen (2008)  
A tri-species *Aspergillus* microarray: Comparative transcriptomics of three *Aspergillus* species. *Proc. Nat. Acad. Sci.* 105:4387-4392
- (241) W. Vongsangnak; P. Olsen; K. Hansen; S. Krosgaard; J. Nielsen (2008)  
Improved annotation through genome-scale metabolic modeling of *Aspergillus oryzae*. *BMC Genomics* 9:245
- (242) A. Fazio; M. C. Jewett; P. Daran-Lapujade; R. Mustacchi; R. Usaite; J. Pronk; C. Workman; J. Nielsen (2008)  
Transcription factor control of growth rate dependent genes in *Saccharomyces cerevisiae*: A three factor design. *BMC Genomics* 9:341
- (243) S. Wattanachaisaerekul; A. E. Lantz; M. L. Nielsen; J. Nielsen (2008)  
Production of the polyketide 6-MSA in yeast engineered for increased malonyl-CoA supply. *Metabolic Eng.* 10:246-254
- (244) I. Nookaew; M. C. Jewett; A. Meecha; C. Thammarongtham; K. Loateng; S. Cheevadhanarak; J. Nielsen; S. Bhumirata (2008)  
The genome-scale metabolic model iIN800 of *Saccharomyces cerevisiae* and its validation: a scaffold to query lipid metabolism. *BMC Sys. Biol.* 2:71
- (245) F. Pizarro; M. C. Jewett; J. Nielsen; E. Agosin (2008)  
Growth temperature exerts a differential physiological and transcriptional response in laboratory and wine strains of *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* 74:6358-6368
- (246) I. Borodina; J. Siebring; J. Zhang; C. P. Smith; G. van Keulen; L. Dijkhuizen; J. Nielsen (2008)

- Antibiotic overproduction in *Streptomyces coelicolor* A3(2) mediated by phosphofructokinase deletion. *J. Biol. Chem.* **283**:25186-25199
- (247) J. Højer-Pedersen; J. Smedsgaard; J. Nielsen (2008)  
The yeast metabolome addressed by electrospray ionization mass spectrometry: Initiation of a mass spectral library and its application for metabolic footprinting by direct infusion mass spectrometry. *Metabolomics* **4**:393-405
- (248) M. J. Hørgård; N. et al.; J. Nielsen; D. B. Kell (2008)  
A consensus yeast metabolic network obtained from a community approach to systems biology. *Nature Biotechnol.* **26**:1155-1160
- (249) M. Rocha; P. Maia; R. Mendes; J. P. Pinto; E. C. Ferreira; K. Patil; J. Nielsen; I. Rocha (2008)  
Natural computation meta-heuristics for the in silico optimization of microbial strains. *BMC Bioinformatics* **9**:499
- (250) J. Maury; M. A. Asadollahi; K. Møller; M. Schalk; A. Clark; L. R. Formenti; J. Nielsen (2008)  
Reconstruction of a bacterial isoprenoid biosynthetic pathway in *Saccharomyces cerevisiae*. *FEBS Let.* **582**:4032-4038
- (251) G. Panagiotou; M. R. Andersen; T. Grotkjær; T. B. Regueira; G. Hofmann; J. Nielsen; L. Olsson (2008)  
Systems analysis unfolds the relationship between the phosphoketolase pathway and growth in *Aspergillus nidulans*. *PLoS One* **3**:e2847
- 2009**
- (252) J. R. Wortman et al. (2009)  
The 2008 update of the *Aspergillus nidulans* genome annotation: A community effort. *Fungal Gen. Biol.* **46**:S2-S13
- (253) K. R. Kjeldsen; J. Nielsen (2008)  
In silico genome-scale reconstruction and validation of the *Corynebacterium glutamicum* metabolic network. *Biotechnol. Bioeng.* **102**:583-597
- (254) P. M. Coutinho et al. (2009)  
Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. *Fungal Gen. Biol.* **46**:S161-S169
- (255) G. Hofmann; A. Diano; J. Nielsen (2009)  
Recombinant bacterial haemoglobin alters metabolism of *Aspergillus niger*. *Met. Eng.* **11**:8-12
- (256) J. Thykær; K. Ruckenstein; H. Noorman; J. Nielsen (2009)  
Disruption of the NADPH-dependent Glutamate Dehydrogenase Affects the Morphology of Two Industrial Strains of *Penicillium chrysogenum*. *J. Biotechnol.* **139**:280-282
- (257) T. S. Christensen; A. P. Oliveira; J. Nielsen (2009)  
Reconstruction and logical modeling of glucose repression signaling pathways in *Saccharomyces cerevisiae*. *BMC Systems Biol.* **3**:7
- (258) S. Meijer; J. Otero; M. R. Andersen; L. Olsson; J. Nielsen (2009)  
Overexpression of isocitrate lyase – glyoxylate bypass influence on metabolism in *Aspergillus niger*. *Metabolic Eng.* **11**:107-116
- (259) D. Cimini; K. R. Patil; C. Shiraldi; J. Nielsen (2009)  
Global transcriptional response of *Saccharomyces cerevisiae* to the deletion of SDH3. *BMC Systems Biol.* **3**:17
- (260) J. F. Moxley; M. C. Jewett; M. R. Antoniewicz; S. G. Villas-Boas; H. Alper; R. T. Wheeler; L. Tong; A. G. Hinnebusch; T. Ideker; J. Nielsen; G. Stephanopoulos (2009)  
Linking transcriptional regulation and high resolution metabolic fluxes in yeast modulated by the global regulator Gcn4p. *Proc. Nat. Acad. Sci.* **106**:6477-6482
- (261) G. Panagiotou; M. R. Andersen; T. Grotkjær; T. B. Regueira; J. Nielsen; L. Olsson (2009)  
Studies of the production of fungal polyketides in *Aspergillus nidulans* by using systems biology tools. *Appl. Environ. Microbiol.* **75**:2212-2220
- (262) S. Meier; M. L. Nielsen; L. Olsson; J. Nielsen (2009)  
Gene deletion of cytosolic ATP-Citrate lyase leads to altered organic acid production in *Aspergillus niger*. *J. Ind. Microbiol. Biotechnol.*, in press
- (263) M. R. Andersen; L. Lehmann; J. Nielsen (2009)  
Systemic analysis of the response of *Aspergillus niger* to ambient pH. *Genome Biol.* **10**:R47

## Reviews in international journals with peer review

- (1) J. Nielsen, J. Villadsen (1992)  
Modelling of microbial kinetics. *Chem. Eng. Sci.* **47**:4225-4270
- (2) J. Nielsen (1996)  
Modelling the morphology of filamentous microorganisms. *Trends Biotechnol.* **14**:438-443
- (3) L. Olsson; J. Nielsen (1997)  
On-line and in-situ monitoring of biomass in submerged cultivations. *Trends Biotechnol.* **15**:517-522
- (4) C. Klein; L. Olsson; J. Nielsen (1998)  
Glucose control in *Saccharomyces cerevisiae*. The role of *MIG1* on metabolic functions. *Microbiol.* **144**:13-24
- (5) L. Olsson; U. Schulze; J. Nielsen (1998)  
On-line bioprocess monitoring - an academic discipline or an industrial tool? *Trends Anal. Chem.* **17**:88-95

- (6) C. M. Henriksen, J. Nielsen, J. Villadsen (1998)  
Cyclization of  $\alpha$ -amino adipic acid into the  $\delta$ -lactam 6-oxo-piperidine-2-carboxylic acid by *Penicillium chrysogenum*. *J. Antibio.* **51**:99-106
- (7) J. Nielsen (1998)  
The role of metabolic engineering in the production of secondary metabolites. *Curr. Opin. Microbiol.* **1**:330-336
- (8) S. Ostergaard, L. Olsson, J. Nielsen (2000)  
Metabolic engineering of *Saccharomyces cerevisiae*. *Microb. Mol. Biol. Rev.* **64**:34-50
- (9) A. K. Gombert, J. Nielsen (2000)  
Mathematical modelling of metabolism. *Curr. Opin. Biotechnol.* **11**:180-186
- (10) J. Nielsen (2001)  
Metabolic engineering. *Appl. Microbiol. Biotechnol.* **55**:263-283
- (11) J. Zaldívar, J. Nielsen, L. Olsson (2001)  
Metabolic engineering of biocatalysts for fuel ethanol production. *Appl. Microbiol. Biotechnol.* **56**:17-34
- (12) J. Thykær, J. Nielsen (2003)  
Metabolic engineering of  $\beta$ -lactam production. *Metabolic Eng.* **5**:56-69
- (13) G. Hofmann, M. McIntyre, J. Nielsen (2003)  
Fungal genomics beyond *Saccharomyces cerevisiae*. *Curr. Op. Biotechnol.* **14**:226-231
- (14) K. Patil, M. Åkesson, J. Nielsen (2004)  
Use of genome scale microbial models for metabolic engineering. *Curr. Op. Biotechnol.* **15**:64-69
- (15) I. Borodina, J. Nielsen (2005)  
From genomes to *in silico* cells via metabolic networks. *Curr. Op. Biotechnol.* **16**:1-6
- (16) S. G. Villas-Bôas, S. Mas, M. Åkesson, J. Smedsgaard, J. Nielsen (2005)  
Mass spectrometry in metabolome analysis. *Mass Spec. Rev.* **24**:613-646
- (17) J. Nielsen, S. Oliver (2005)  
The next wave in metabolome analysis. *Trends Biotechnol.* **23**:544-546
- (18) M. C. Jewett, G. Hofmann, J. Nielsen (2006)  
Fungal metabolite analysis in genomics and phenomics. *Curr. Op. Biotechnol.* **17**:1-7
- (19) S. Rokem, A. Eliasson, J. Nielsen (2007)  
Systems biology of secondary metabolism in microorganisms. *Nat. Prod. Rep.* **24**:1262-1287
- (20) J. Nielsen, M. Jewett (2008)  
Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **8**:122-131
- (21) V. Mapelli, L. Olsson, J. Nielsen (2008)  
Metabolic footprinting in microbiology: methods and applications in functional genomics and biotechnology. *Trends Biotechnol.* **26**:490-497
- (22) D. Petranovic, J. Nielsen (2008)  
Can yeast systems biology contribute to the understanding of human disease? *Trends Biotechnol.* **26**:584-590
- (23) M. R. Andersen, J. Nielsen (2009)  
Current status of systems biology in Aspergilli. *Fun. Gen. Biol.* **46**:S180-S190
- (24) G. Panagiotou, J. Nielsen (2009)  
Nutritional systems biology: Definitions and approaches. *Annual Rev. Nutr.*, in press

## Commentaries and Short Papers

- (1) B. Hahn-Hägerdal, K. van Dam, I. Spencer-Martins, J. Nielsen (1998)  
Le dédale du métabolisme. *Biofutur* **184**:81-82
- (2) J. Thykær, B. Christensen, J. Nielsen (2001)  
The application of metabolic network analysis in metabolic engineering. *BIOforum International* **5**:228-231
- (3) J. M. Otero, L. Olsson, J. Nielsen (2007)  
Industrial biotech meets systems biology. *Genetic Eng. News* **1**:28-31
- (4) J. Nielsen (2007)  
Principles of optimal metabolic network operation. *Mol. Sys. Biol.* **3**:126
- (5) G. Vemuri, J. Nielsen (2008)  
Systems biology: Is the hope worth the hype. *SIM News* **58**:178-188
- (6) J. Nielsen (2008)  
Vintage paper Introduction. *Biotechnol. Bioeng.* **103**:1

## Books

- (1) M. Y. Andersen *et al.* (1991)  
Bioreaktionsteknik. Metoder til fortolkning af data fra bioreaktorer. Akademiet for Tekniske Videnskaber, Lyngby
- (2) J. Nielsen, J. Villadsen (1994)  
Bioreaction Engineering Principles. Plenum Press, New York
- (3) J. Nielsen (1997)

- (4) Physiological Engineering Aspects of *Penicillium chrysogenum*. World Scientific Publishing Co., Singapore
- (5) G. Stephanopoulos; A. Aristodou; J. Nielsen (1998)  
Metabolic Engineering. Academic Press, San Diego
- (6) J. Nielsen (2001) (editor)  
Metabolic Engineering. Advances in Biochemical Engineering/Biotechnology, Vol. 73, Springer Verlag, Heidelberg
- (7) J. Nielsen; J. Villadsen; G. Liden (2003)  
Bioreaction Engineering Principles, 2. ed., Kluwer Plenum, New York
- (8) J. Nielsen (2006) (editor)  
Frontiers in Biotechnology. Advances in Biochemical Engineering/Biotechnology, Vol. 100, Springer Verlag, Heidelberg
- (9) S. G. Villas-Boas; U. Roessner; M. A. E. Hansen; J. Smedsgaard; J. Nielsen (2007)  
Metabolome Analysis. An Introduction, Wiley, Hoboken
- (10) J. Nielsen; M. C. Jewett (2008) (editors)  
Metabolomics. Topics in Current Genetics, Vol. 18, Springer Verlag, Heidelberg

## Contribution in Books

- (1) J. Nielsen (1992)  
Modelling the growth of filamentous fungi. Adv. Biochem. Eng./Biotechnol. 46:187-223
- (2) J. Nielsen; J. Villadsen (1993)  
Bioreactors: Description and modelling. pp. 77-104, Chap. 5 in Biotechnology Vol. 3 (2. ed.). Eds. H.-J. Rehm; G. Reed (volume editor G. Stephanopoulos), VCR Verlag
- (3) J. Nielsen; M. Carlsen (1996)  
Fungal Pellets. pp. 273-293, Chap. 13. In: Immobilised living cell systems: Modelling and experimental methods. Eds. G. Baron, R. Willaert; L. De Backer, John Wiley & Sons, Chichester
- (4) P. Krahbenz; J. Nielsen (1998)  
Modeling the mycelium morphology of *Penicillium* species in submerged cultures. Adv. Biochem. Eng./Biotechnol. 60:125-152
- (5) A. B. Spohr, T. Agger, M. Carlsen; J. Nielsen (1998)  
Quantitative morphology of filamentous microorganisms. pp. 373-410, Chap. 14. In: Digital analysis of Microbes. Eds. M. H. F. Wilkinson; F. Schut, John Wiley & Sons
- (6) J. Nielsen (1999)  
Fermentation monitoring, design and optimization. pp. 1147-1157. In: The Encyclopedia of Bioprocess Technology: Fermentation, Biocatalysis & Bioseparations. Eds. M. C. Flickinger; S. W. Drew, John Wiley & Sons
- (7) J. Nielsen (1999)  
Fermentation kinetics. pp. 69-119, Chap. 4. In: Modern Concepts in Fermentation Biotechnology. Eds. M. El-Mansi, Bryce, Taylor and Francis, London
- (8) B. Christensen; J. Nielsen (1999)  
Metabolic network analysis – powerful tool in metabolic engineering. Adv. Biochem. Eng./Biotechnol. 66:209-231
- (9) J. Nielsen (2001)  
Microbial Process Kinetics. pp. 127-149, Chap. 6. In: Basic Biotechnology. 2. Edition. Eds. C. Ratledge, B. Kristiansen, Cambridge University Press, Cambridge
- (10) J. Nielsen (2001)  
Metabolic Turnover. In: Encyclopaedia of Life Sciences. Nature Publishing Group, London ([www.els.net](http://www.els.net))
- (11) T. Agger; J. Nielsen (2001)  
Mathematical modelling of microbial processes – Motivation and means. pp. 61-75. In: Engineering and Manufacturing for Biotechnology. Eds. M. Hofman, P. Thonart, Kluwer Academic Publishers, Dordrecht
- (12) S. Ostergaard; L. Olsson; J. Nielsen (2001)  
Metabolic Pathway Analysis of *Saccharomyces cerevisiae*. pp. 75-85. In: Applied Microbiology. Focus on Biotechnology. Vol. 2. Eds. M. Hofman, P. Thonart, Kluwer Academic Publishers, Dordrecht
- (13) M. McIntyre; C. Müller; J. Dynesen; J. Nielsen (2001)  
Metabolic engineering of the morphology of *Aspergillus*. Adv. Biochem. Eng./Biotechnol. 73:103-128
- (14) J. Nielsen (2002)  
Metabolic Engineering. pp. 391-406. In: Encyclopedia of Physical Science and Technology, Vol. 9. Academic Press, San Diego
- (15) M. M. dos Santos; M. Åkesson; J. Nielsen (2003)  
Metabolic Flux Analysis in the post genomics area. pp. 89-105. In: Metabolic Engineering in the Post Genomic Era. Eds. B. Kholodenko; H. Westerhoff, Horizon Scientific Press, Norfolk
- (16) C. Bro; B. Regenberg; J. Nielsen (2003)  
Yeast Functional Genomics and Metabolic Engineering; Past, present and future. pp. 331-360. In: Functional Genetics of Industrial Yeast. Ed. H. de Winder, Springer, Heidelberg
- (17) A. Gombert; J. Nielsen (2003)  
Quantification of metabolic fluxes. In: Encyclopaedia of Life Sciences. Nature Publishing Group, London ([www.els.net](http://www.els.net))
- (18) N. Gunnarsson, A. Eliasson, J. Nielsen (2004)



- Control of fluxes towards antibiotics and the role of primary metabolism in production of antibiotics. *Adv. Biochem. Eng./Biotechnol.* **88**: 137-178
- (19) **J. Nielsen**, A. Eliasson (2005)  
From glucose to antibiotics: What controls the fluxes? pp. 195-214. In: *Biocombinatorial Approaches for Drug Finding*. Ed. W. Wohlleben, T. Spelling, B. Müller-Tiemann, Springer, Heidelberg
  - (20) H. David, **J. Nielsen** (2005)  
Modelling of fungal metabolism. Pp. 195-214. In: *Metabolome Analysis. Strategies for Systems Biology*. Ed. S. Vaidyanathan, G. G. Harrigan, R. Goodacre, Springer, New York
  - (21) J. Maury, M. A. Asadollahi, K. Möller, A. Clark, **J. Nielsen** (2005)  
Microbial isoprenoid production: An example of green chemistry through metabolic engineering. *Adv. Biochem. Eng./Biotechnol.* **100**: 19-51
  - (22) J. Højer-Pedersen, J. Smedsgaard, **J. Nielsen** (2006)  
Elucidating the mode-of-action of compounds from metabolite profiling studies, Pp. 103-130. In: *Systems Biological Approaches in Infectious Diseases*. Progress in Drug Research Vol. 64. Ed. H. I. Boshoff, C. E. Barry III, Birkhäuser, Berlin
  - (23) **J. Nielsen** (2006)  
Microbial Process Kinetics, pp. 155-180, Chap. 6. In: *Basic Biotechnology*. 3. Edition. Eds. C. Ratledge, B. Kristiansen, Cambridge University Press, Cambridge
  - (24) **J. Nielsen** (2006)  
Fermentation Kinetics, pp. 69-120. Chap. 4. In: *Fermentation Microbiology and Biotechnology*. Eds. E. M. T. El-Mansi, C. F. A. Bryce, A. L. Demain, A. R. Allman, Taylor and Francis, London
  - (25) I. Rocha, J. Förster, **J. Nielsen** (2007)  
Design and application of genome-scale reconstructed metabolic models, Pp. 409-431. In: *Methods in Molecular Biology*, Vol. 416: *Gene Essentiality*. Ed. S. Y. Gerdes, A. L. Osterman, Humana Press Inc., Totowa, USA
  - (26) W. de Jongh, **J. Nielsen** (2007)  
Filamentous fungi as cell factories for metabolite production, Pp. 163-182. In: *Food Mycology. A multifaceted approach to fungi and food*. Ed. J. Dijksterhuis, R. A. Samson, CRC Press, Boca Raton, USA
  - (27) A. P. Oliveira, M. C. Jewett, **J. Nielsen** (2007)  
From gene expression to metabolic fluxes, Pp. 37-66. In: *Introduction to Systems Biology*. Ed. S. Choi, Humana Press, Totowa, USA
  - (28) M. C. Jewett, **J. Nielsen** (2008)  
The role of Metabolomics in systems biology. Pp. 1-10. In: *Metabolomics. Topics in Current Genetics*. Eds. J. Nielsen and M. C. Jewett, Vol. 18, Springer Verlag, Heidelberg
  - (29) M. C. Jewett, M. A. E. Hansen, **J. Nielsen** (2008)  
Data acquisition, analysis, and mining: Integrative tools for discerning metabolic function in *Saccharomyces cerevisiae*. Pp. 159-188. In: *Metabolomics. Topics in Current Genetics*. Eds. J. Nielsen and M. C. Jewett, Vol. 18, Springer Verlag, Heidelberg

## Papers in Conference Proceedings

- (1) **J. Nielsen**; K. Nikolajsen; J. Villadsen (1989)  
Computer controlled system for on-line monitoring of a fermentation process. pp. 53-57. In *Computer Applications in Fermentation Technology. Modelling and Control of Biotechnological Processes*, eds. N. M. Fish; R. I. Fox; N. F. Thornhill, Elsevier Applied Science, London
- (2) J. Villadsen; **J. Nielsen** (1990)  
Modelling of fermentation kinetics. pp. 259-266. Proc. V<sup>th</sup> European Congress on Biotechnology, Lyngby
- (3) **J. Nielsen** (1991)  
Application of structured fermentation models. pp. 843-848. Proc. European Simulation Multiconference, Copenhagen
- (4) S. Benhin; **J. Nielsen**; J. Villadsen (1992)  
Transmembrane transport systems, studied on actively growing cells. pp. 351-355. In *Harnessing Biotechnology for the 21st Century* (Proc. IX<sup>th</sup> International Biotechnology Symposium, Crystal City), eds. M. R. Ladisch; A. Rose, **J. Nielsen** (1993)
- (5) Modelling of filamentous microorganisms. pp. 45-52. In *Computer Applications in Fermentation Technology. Modelling and Control of Biotechnological Processes*, eds. M. N. Karim; G. Stephanopoulos, Pergamon Press
- (6) H. S. Jørgensen; H. Møllgaard; **J. Nielsen**; J. Villadsen (1993)  
Identification of rate controlling enzymes in the metabolic pathway of penicillin in a high yielding strain of *P. chrysogenum*. pp. 251-254. In *Computer Applications in Fermentation Technology. Modelling and Control of Biotechnological Processes*, eds. M. N. Karim; G. Stephanopoulos, Pergamon Press
- (7) **J. Nielsen** (1993)  
Modelling the morphology of filamentous fungi. pp. 131-142. Proc. Bioreactor Performance, eds. U. Mortensen; H. J. Noorman, Elsingore
- (8) S. Frandsen; **J. Nielsen**; J. Villadsen (1993)  
Application of regimen analysis of yeast fermentation for down-scaling. pp. 171-179. Proc. Bioreactor Performance, eds. U. Mortensen; H. J. Noorman, Elsingore

- (9) S. Frandsen; **J. Nielsen**; J. Villadsen (1993)  
Dynamics of *Saccharomyces cerevisiae* in continuous culture. pp. 887-890. In Progress in Biotechnology, Vol. 9. Proc. VI<sup>th</sup> European Congress on Biotechnology, Florence. Elsevier
- (10) A. G. Pedersen; **J. Nielsen**; J. Villadsen (1993)  
Characterization of bioreactors using isotope techniques. pp. 931-934. In Progress in Biotechnology, Vol. 9. Proc. VI<sup>th</sup> European Congress on Biotechnology, Florence. Elsevier
- (11) **J. Nielsen** (1994)  
Physiological Engineering - Towards a new science. pp. 30-38. In Proc. The 1994 IChemE Research Event, London, Vol. I
- (12) M. Carlsen; A. Spohr; R. Mørkeberg; **J. Nielsen**; J. Villadsen (1994)  
Growth and protein formation of recombinant *Aspergillus*: Utility of morphological characterization by image analysis. pp. 197-202. In Proc. Advances in Bioprocess Engineering, Cuernavaca. Kluwer Academic Publishers
- (13) **J. Nielsen** (1994)  
Physiological aspects of *Penicillium chrysogenum*. pp. 15-18. Proc. Modelling of filamentous fungi, Otecek.
- (14) G. Lidén; U. Schulze; **J. Nielsen**; J. Villadsen (1995)  
On the optimization of ethanol production in nitrogen limited yeast fermentations. In Proc. VII<sup>th</sup> European Congress on Biotechnology, Nice
- (15) **J. Nielsen**, H. Jørgensen (1995)  
A kinetic model for the penicillin biosynthetic pathway in *Penicillium chrysogenum*. pp. 136-141. In Proc. Computer Application in Biotechnology 6, Garmisch-Partenkirchen
- (16) P. N. Pissarra; M. J. Bazin; K. Schmidt; **J. Nielsen** (1995)  
A user-friendly and interactive methodology for simulation and non-linear parameter estimation of biotechnological process models using SIMULINK. Proc. Nordic Matlab Conference '95, Stockholm
- (17) K. Schmidt; P. N. Pissarra; **J. Nielsen** (1995)  
The evolutionary strategy: A simple and straightforward optimization algorithm for parameter estimation in structured biochemical models. Proc. Nordic Matlab Conference '95, Stockholm
- (18) **J. Nielsen** (1995)  
Metabolic control analysis of the penicillin biosynthetic pathway. Proc. National Biotechnology and Bioengineering Conference, Ixtapa, Mexico
- (19) U. Schulze; T. L. Nissen; **J. Nielsen**; J. Villadsen (1996)  
Application of Metabolic Flux Analysis in physiological studies. Proc. V<sup>th</sup> World Congress on Chemical Engineering, San Diego, USA
- (20) P. de N. Pissarra; **J. Nielsen** (1996)  
Kinetic and thermodynamic analysis of the penicillin biosynthetic pathway. Proc. V<sup>th</sup> World Congress on Chemical Engineering, San Diego, USA
- (21) D. L. Bogle et al. (1996)  
Process synthesis, design and simulation of integrated biochemical processes. Proc. V<sup>th</sup> World Congress on Chemical Engineering, San Diego, USA
- (22) **J. Nielsen**; I. Spencer Martins (1997)  
Metabolic Engineering. Proc. European Cell Factory Conference, Lund, Sweden
- (23) S. Østergaard; L. Olsson; **J. Nielsen** (1998)  
Metabolic control analysis of the Leloir pathway in *Saccharomyces cerevisiae*. BioThermoKinetics in the post genomic era. Proc. 8<sup>th</sup> BTK Meeting, Fiskebäckskil, Sweden
- (24) H. A.æ Theilgaard; **J. Nielsen** (1998)  
The influence of the LLD-ACV-bisACV ration on a MCA of the penicillin biosynthetic pathway in *Penicillium chrysogenum*. BioThermoKinetics in the post genomic era. Proc. 8<sup>th</sup> BTK Meeting, Fiskebäckskil, Sweden
- (25) **J. Nielsen** (1998)  
Mathematical modelling of biochemical pathways. BioThermoKinetics in the post genomic era. Proc. 8<sup>th</sup> BTK Meeting, Fiskebäckskil, Sweden

## Other Publications

- (1) **J. Nielsen** (1990)  
Post doc på Universitét Hannover. Dansk Kemi 5:178-179
- (2) **J. Nielsen** (1990)  
On-line måling af biomasse i bioreaktorer. Dansk Kemi 8:260-265
- (3) **J. Nielsen**; J. Gram; L. Foldager (1990)  
Bioreaktionsteknik: Kendte metoder anvendt på nye processer. Dansk Kemi 12:430-431
- (4) **J. Nielsen** (1993)  
Physiological Engineering. Dansk Kemi 11:14-17
- (5) **J. Nielsen**; A. G. Pedersen; M. Bundgaard-Nielsen; H. Andersen (1994)  
Anvendelse af radioaktive isotoper til karakterisering af bioreaktorer. Dansk Kemi 4:28-31
- (6) **J. Nielsen** (1995)  
Doktorafhandling om industriel penicillinproduktion. Dansk Kemi 11:20-23
- (7) **J. Nielsen** (1995)

- Penicillinets historie. Dansk Kemi 12:24-28
- (8) U. Schulze; **J. Nielsen** (1997)  
Massachusetts Institute of Technology. Dansk Kemi
- (9) T. Agger; **J. Nielsen** (1999)  
Optimering af enzymproduktion med skimmelsvampe. Dansk Kemi 1:14-15
- (10) B. Christensen; **J. Nielsen** (1999)  
Metabolsk flux analyse – en kvantitativ beskrivelse af mikroorganismeres primære metabolisme. Dansk Kemi 1:16-18
- (11) **J. Nielsen** (1999)  
Functional genomics og dens rolle i udviklingen af fremtidige bioteknologiske processer. Dansk Kemi 1:20-27
- (12) C. Bro; B. Regenber; **J. Nielsen** (2001)  
DNA Arrays. Dansk Kemi 1:18-20
- (13) M. R. Andersen; M. L. Nielsen; **J. Nielsen** (2006)  
Genomet – systembiologiens ryggrad. Dansk Kemi 3:15-17
- (14) J. Højer-Pedersen; J. Smedsgaard; **J. Nielsen** (2006)  
Metabolomet: Et indirekte produkt af genomet. Dansk Kemi 3:18-20

**FINAL PROGRAM**

**BIOCHEMICAL ENGINEERING XI:  
MOLECULAR DIVERSITY IN  
DISCOVERY AND BIOPROCESSING**

July 25 - 30, 1999  
Marriott University Park Hotel  
480 Wakara Way  
Salt Lake City, Utah 84108  
1-801-581-1000 - Fax: 1-801-584-3321

**Conference Chair:**

George Georgiou  
University of Texas

**Conference Co-Chair:**

Steven Lee  
Merck & Co, Inc.

**United Engineering Foundation Inc.**

Three Park Avenue, 27th Floor  
New York, NY 10016-5902  
Phone: 1-212-591-7836 Fax: 1-212-591-7441  
E-mail: engfnd@aol.com Website: www.engfnd.org

**Biochemical Engineering XI Advisory Committee**

**Juan A. Asenjo**  
*University of Chile*

**Alain Bernard**

*Serono Pharmaceutical Research Institute*

**Arindam Bose**

*Pfizer, Inc.*

**Douglas C. Cameron**

*Cargill*

**Gopal K. Chotani**

*Genencor International*

**Stephen W. Drew**

*Merck & Co., Inc.*

**Wei Shou Hu**

*University of Minnesota*

**Robert Kelly**

*North Carolina State University*

**Allen Laskin**

*Laskin/Lawrence Associates*

**Sang Yup Lee**

*KAIST*

**Vasantha Nagaragan**

*Dupont*

**David Naveh**  
*Bayer Biotechnology*  
**Jens Nielsen**  
*Technical University of Denmark*  
**E. Terry Papoutsakis**  
*Northwestern University*  
**Greg Stephanopoulos**  
*Massachusetts Institute of Technology*  
**James R. Swartz**  
*Stanford University*  
**K. Dane Witttrup**  
*University of Illinois*  
**Toshiomi Yoshida**  
*Osaka University*  
**Dane W. Zabriskie**  
*SmithKline Beecham Laboratories*

**Financial Sponsors of Biochemical Engineering XI Abbott Laboratories**

**Amgen, Inc.**  
**Bayer Corporation**  
**Biogen**  
**Boehringer Ingelheim**  
**Bristol-Myers Squibb**  
**Covance Bio**  
**E.I. Du Pont de Nemours and Company**  
**Genencor International**  
**Genentech, Inc.**  
**Gist-Brocades**  
**Lilly Research Laboratories (Eli Lilly & Co.)**  
**Merck & Co., Inc.**  
**Novo Nordisk A/S**  
**National Science Foundation, U.S.A.**  
**Pfizer, Inc. The Proctor and Gamble Company**  
**Schering-Plough Research Institute**  
**Serono Pharmaceutical Research Institute**  
**U.S. Department of Army (Army Research Office)**  
**The Whitaker Foundation**

**Poster Session A**

**Poster Session B**

**Sunday, July 25, 1999**

12:00 noon - 2:30 p.m. Conference Registration (Ballroom Atrium)

2:30 p.m. - 3:00 p.m. **Welcoming Remarks** (Ballroom I)

George Georgiou, University of Texas at Austin

Steven Lee, Merck & Co., Inc.

Allen Laskin, United Engineering Foundation

**PLENARY SPEAKERS:**

3:00 p.m. - 4:00 p.m. *Whole-Genome RNA and Protein Regulatory Networks in E-Coli and S. Cerevisiae*

G. Church  
Harvard University

4:00 p.m. - 5:00 p.m. *The Chemical Industry in the 21st Century: Role of Biotechnology*

J. Miller  
CTO, DuPont

5:00 p.m. - 5:30 p.m. Break (Ballroom Atrium)

5:30 p.m. - 6:30 p.m. *Using New Gene Targets in Drug Discovery with Effective Throughput Screening Systems*

P. Fernandes  
CEO, Small Molecule Therapeutics

6:30 p.m. - 7:30 p.m. *Methods for Mining DNA Microarray Data*

G. Stephanopoulos  
Department of Chemical Engineering, Massachusetts Institute of Technology, USA

7:30 p.m. - 9:00 p.m. Dinner (Ballroom II & III)

9:00 p.m. - 10:30 p.m. Opening Reception (Ballroom Atrium)

#### **Monday, July 26, 1999**

7:00 a.m. - 8:15 a.m. Breakfast (Ballroom II & III)

**SESSION (a): EXPLOITING MICROBIAL DIVERSITY** (Ballroom I)

**Session Chairs:** R. Kelly, North Carolina State University  
V. Nagarajan, DuPont

8:15 a.m. - 8:20 a.m. **Introduction:** Session Chairs

8:20 a.m. - 9:00 a.m. *Gene Acquisition in Bacterial Genome Evolution*

J. Roth  
University of Utah

9:00 a.m. - 9:30 a.m. *How Bacteria Talk to Each Other: Quorum Sensing in Escherichia Coli, Salmonella Typhimurium and Vibrio Harveyi*

L. Bassler  
Princeton University

9:30 a.m. - 10:00 a.m. *Predicting the Consequences of Genome Reorganization*

I. Molineux  
University of Texas at Austin

10:00 a.m. - 10:30 a.m. Coffee Break (Ballroom Atrium)

10:30 a.m. - 11:00 a.m. ***Hyperthermophilic Genome Sequences: So Many Interesting Enzymes, So Little Time***

R. Kelly

North Carolina State University

11:00 a.m. - 11:30 a.m. ***Purification, Characterization and Process Considerations of Cryophilic Proteases of Marine Origin***

J. Asenjo

University of Chile

11:30 a.m. - 12:00 noon ***Harvesting Biomolecules for the Tree of Life***

E. Mathur

Diversa Corporation

12:00 noon - 12:30 p.m. ***Genetic and Biochemical Diversity of Bacterial in Commercial Wastewater Bioreactors***

V. Nagarajan

DuPont Company

12:30 p.m. - 1:45 p.m. Lunch (Ballroom II & III)

1:45 p.m. - 4:00 p.m. *Ad hoc* Sessions and/or free time

**SESSION (b): NOVEL HIGH THROUGHPUT TECHNOLOGIES IN DISCOVERY** (Ballroom I)

**Session Chairs:** G. Georgiou, University of Texas at Austin

J. Chalmers, Ohio State University

4:00 p.m. - 4:05 p.m. **Introduction:** Session Chairs

4:05 p.m. - 4:35 p.m. ***Non-Invasive Monitoring of Pathogens, Tumor Cells and Gene Expression in a Living Animal***

P. Contag

CEO, Xenogen Corp.

4:35 p.m. - 5:05 p.m. ***Directed Enzyme Evolution on Bacterial Surfaces***

B.L. Iverson

University of Texas at Austin

5:05 p.m. - 5:35 p.m. ***Lead Identification from Encoded Synthetic Combinatorial Libraries Via a Self-Selective, Positive Feedback Reporter System***

G. Lalonde

Affymax Research Institute

5:35 p.m. - 6:00 p.m. Coffee Break (Ballroom Atrium)

6:00- p.m. - 6:30 p.m. ***Transcript Profile and Proteomics-Based Views of Biological Models for Pharmaceutical Development***

J. Seilhamer

Incyte Pharmaceuticals, Inc.

6:30 p.m. - 7:00 p.m. *Microfabricated Chemical Analyses Systems*

Mark Burns

University of Michigan

7:00 p.m. - 7:30 p.m. *Exploitation of Immunomagnetic Labeling: Rapid Cell Sorting and Fractionation*

J. Chalmers

Ohio-State University, USA

7:30 p.m. - 9:00 p.m. Dinner (Ballroom II & III)

9:15 p.m. - 10:30 p.m. Poster Session A and Social Hour (Ballroom Atrium)

**Tuesday, July 27, 1999**

7:00 a.m. - 8:25 a.m. Breakfast (Ballroom II & III)

**SESSION (C) : DISCOVERY AND DESIGN OF MACROMOLECULES** (Ballroom I)

**Session Chairs:** B.L. Iverson, University of Texas at Austin

V. Schellenberger, Genencor International, USA

8:25 a.m. - 8:30 a.m. **Introduction:** Session Chairs

8:30 a.m. - 9:00 a.m. *Evolution of Ribozymes, Proteins and Pathways*

A. Ellington

University of Texas at Austin

9:00 a.m. - 9:30 a.m. *Directed Evolution of Subtilisin*

V. Schellenberger

Genencor International, USA

9:30 a.m. - 10:00 a.m. *Directed Evolution of Protein Recognition. Stability and Expression by Yeast Surface Display*

K.D. Wittrup

University Illinois

10:00 a.m. - 10:30 a.m. Coffee Break (Ballroom Atrium)

10:30 a.m. - 11:00 a.m. *Creation of New Cytotoxic Antitumor Antibodies by Glycosylation Engineering*

J.E. Bailey

Institute of Biotechnology, ETH Zurich, Switzerland

11:00 a.m. - 11:30 a.m. *Molecular Breeding of Genes, Pathways and Genomes by DNA Shuffling*

W.P.C. Stemmer

Maxygen, Inc., USA

11:30 a.m. - 12:00 noon Discussion



12:00 noon - 1:30 p.m. Lunch (Ballroom II & III)

1:45 p.m. - 3:00 p.m. *Ad hoc* Sessions and/or free time

**SESSION (d): BIOINFORMATICS (Ballroom I)**

**Session Chairs:** J. Nielsen, Technical University Denmark, Denmark  
B.O. Palsson, University of California at San Diego

3:15 p.m. - 3:45 p.m. *Definition of the E. Coli Metabolic Genotype: Basic Concepts, Scientific and Applied Uses*

B.O. Palsson  
University of California at San Diego

3:45 p.m. - 4:15 p.m. *Strategies for Prediction of Orphan Protein Function*

S. Brunak  
Technical University of Denmark

4:15 p.m. - 4:45 p.m. *Pathmap; A New Tool for Visualization of Gene Expression Data*

John Rogers  
Parke-Davis Pharmaceutical Research

4:45 p.m. - 5:15 p.m. *Bioinformatics in the Postgenomic Era*

S. Subramanian  
University of Illinois UC

5:15 p.m. - 5:45 p.m. Coffee Break (Ballroom Atrium)

**SESSION (e): DISCOVERY AND PRODUCTION OF BIOACTIVE SMALL MOLECULES**

**Session Chairs:** A.E. Barron, Northwestern University  
D.S. Clark, University of California - Berkeley

5:45 p.m. - 6:15 p.m. *Integrating Biocatalysis into the Drug Discovery Pipeline*

D.S. Clark  
University of California - Berkeley

6:15 p.m. - 6:45 p.m. *A Stress Promoter-Based System for Antibiotics Screening*

F. Baneyx  
Department of Chemical Engineering, University of Washington, USA

6:45 p.m. - 7:15 p.m. *The Challenge of Ultra-High Throughput Screen Technology: Implementation of Biological Assays to 3,456 Well Format*

L. Mere  
Aurora Biosciences Corporation

7:15 p.m. - 7:45 p.m. *Peptoids: A High Diversity Family of Synthetic, Protease-Stable Peptide Mimics*

A.E. Barron  
Chemical Engineering Department, Northwestern University, USA

7:45 p.m. - 8:00 p.m. Discussion

8:00 p.m. - 9:30 p.m. Dinner (Ballroom II & III)

9:30 p.m. - 10:30 p.m. Social Hour (Ballroom Atrium)

**Wednesday, July 28, 1999**

7:00 a.m. - 8:15 a.m. Breakfast (Ballroom II & III)

**SESSION (f): HIGH THROUGHPUT BIOPROCESSING (Ballroom I)**

**Session Chairs:** S. Ozturk, Bayer Corporation  
S. Drew, Merck & Co.

8:15 a.m. - 8:20 a.m. **Introduction:** Session Chairs

8:20 a.m. - 8:50 a.m. ***An "Animal" on a Chip: Preclinical Evaluation of Pharmaceuticals***

M. L. Shuler  
School of Chemical Engineering  
Cornell University, USA

8:50 a.m. - 9:20 a.m. ***An Efficient Integrated Approach to Bioprocess Development***

R. Greasham  
Merck & Co.

9:20 a.m. - 9:50 a.m. ***Biocatalytic Synthesis of Chiral Intermediate for Anti-Hypertension Drugs***

R. Patel  
Bristol-Myers Squibb Pharmaceutical Research Institute

9:50 a.m. - 10:20 a.m. Coffee Break (Ballroom Atrium)

10:20 a.m. - 10:50 a.m. ***Transient Gene Expression for Biotech Research***

A. Bernard  
Serono Pharmaceutical Research Institute, S.A., Switzerland

10:50 a.m. - 11:20 a.m. ***High Throughput Bioprocessing: New Directions in Pharmaceutical Drug Development***

S. Ozturk  
Bayer Corporation

11:20 a.m. - 11:50 a.m. ***Protein Production in Transgenic Plants***

V. Paradkar  
Monsanto Company

11:50 a.m. - 12:00 noon Discussion

12:00 noon - 1:30 p.m. Lunch (Ballroom II & III)

1:30 p.m. - 4:00 p.m. **Poster Session B/Ad hoc Sessions and/or free time (Ballroom Atrium)**

**SESSION (g): CONTROLLING PRODUCT HETEROGENEITY (Ballroom I)****Session Chairs:** E.T. Papoutsakis, Northwestern University

M.L. Shuler, Cornell University

4:15 p.m. - 4:20 p.m. **Introduction:** Session Chairs4:20 p.m. - 4:50 p.m. *Genetic and Physiological Manipulation of the Protein Glycosylation Pathway*

P. Jenkins

Lilly Research Laboratories, Eli Lilly &amp; Co.

4:50 p.m. - 5:20 p.m. *Controlling Secretory Processing to Improve Product Quality and Homogeneity*

M.J. Betenbaugh

Department of Chemical Engineering, Johns Hopkins University, USA

5:20 p.m. - 5:50 p.m. *DNA Microarray for Metabolic Engineering*

J.C. Liao

Department of Chemical Engineering, University of California Los Angeles, USA

5:50 p.m. - 6:15 p.m. Coffee Break (Ballroom Atrium)

6:15 p.m. - 6:45 p.m. *The Domain-Bypass Mechanism for Creation of Structural Diversity in the Pikromycin Biosynthetic Gene Cluster*

D. Sherman

University of Minnesota

6:45 p.m. - 7:15 p.m. *Sugars to Plastics, then to Fine Chemical Monomers: Production of Various Enantiomerically Pure @- Hydroxycarboxylic Acids*

S.Y. Lee

Chemical Engineering Department, Korea Advanced Institute of Science and Technology, Korea

7:30 p.m. - 10:00 p.m. Conference Banquet/AMGEN Award and Social Hour (Ballroom II &amp; III)

**AMGEN Award Lecture***Cell Engineering: Understanding and Controlling Receptor Processes*

Douglas A. Lauffenburger

Massachusetts Institute of Technology

**Thursday, July 29, 1999**

7:00 a.m. - 8:15 a.m. Breakfast (Ballroom II &amp; III)

**SESSION (h): DISCOVERY OF BIOLOGICAL MATERIALS (Ballroom I)****Session Chairs:** D. Kaplan, Tufts

S.Y. Lee, KAIST, Korea

8:15 a.m. - 8:20 a.m. **Introduction:** Session Chairs8:20 a.m. - 8:50 a.m. *Bioengineered Microbial Lipopolysaccharides and Glycolipids*

D. Kaplan

Tufts

8:50 a.m. - 9:20 a.m. *Bioactive Biomaterials for Engineering Tissue Healing*  
J.A. Hubbell  
ETH Zürich

9:20 a.m. - 9:50 a.m. *Molecular Weight Control in Bacterial Hyaluronic Acid Production*  
L. Nielsen  
Chemical Engineering Department, University of Queensland, Australia

9:50 a.m. - 10:20 a.m. Coffee Break (Ballroom Atrium)

10:20 a.m. - 10:50 a.m. *S-Layers: Fundamentals and Applications in Molecular Nanotechnology and Biomimetics*  
Prof. U.B. Sleytr  
University of Agricultural Sciences, Vienna, Austria

10:50 a.m. - 11:20 a.m. *Bioinductive Polymer Coatings for Implantable Glucose Biosensors*  
J. Koberstein  
University of Connecticut

11:20 a.m. - 11:50 a.m. *Properties and Biotechnological Applications of the Azotobacter Vinelandii Modular Type Mannuronan C-5 Epimerases*  
S. Valla  
Norwegian University of Science and Technology

11:50 a.m. - 12:00 noon Discussion

12:00 noon - 1:30 p.m. Lunch (Ballroom II & III)

1:30 p.m. - 3:30 p.m. *Ad hoc* Sessions and/or free time

**SESSION (i): PRODUCTION OF COMPLEX PRODUCTS** (Ballroom I)  
**Session Chair:** J.R. Swartz, Stanford University

3:45 p.m. - 3:50 p.m. *Cell Free Protein Synthesis for Discovery and Production*  
J.R. Swartz  
Stanford University

4:20 p.m. - 4:50 p.m. *Production of Viral Vaccines: A Role For Biochemical Engineering*  
D. Robinson  
Merck & Co., USA

4:50 p.m. - 5:20 p.m. *O2 and its Transport in Hematopoietic Life and Death*  
E.T. Papoutsakis  
Northwestern University, USA

5:20 p.m. - 5:45 p.m. Coffee Break (Ballroom Atrium)

5:45 p.m. - 6:15 p.m. *Stem Cell-Based Cellular and Tissue Engineering*

P. Zandstra

University of Toronto, Canada

6:15 p.m. - 6:45 p.m. *The Production of Plasmid DNA for Gene Therapy: The Cell Lysis Step*

A.W. Nienow

Centre for Bioprocess Engineering, School of Chemical Engineering, The University of Birmingham, UK

6:45 p.m. - 7:15 p.m. *Manipulation of the Glycosylation of Recombinant Proteins Produced in CHO Cells by Overexpression of Glycosyltransferase*

L. Krummen

Genentech, Inc., USA

7:15 p.m. - 7:30 p.m. Discussion

7:30 p.m. - 9:00 p.m. Dinner (Ballroom II & III)

9:00 p.m. - 10:00 p.m. Social Hour (Ballroom Atrium)

### **Friday, July 30, 1999**

7:00 a.m. - 8:30 a.m. Breakfast (Ballroom II & III)

8:30 a.m. - 10:00 a.m. Wrap-up Discussion (Ballroom I)

10:00 a.m. - 12:00 noon Free Time

12:00 noon - 1:30 p.m. Lunch (Ballroom II & III)

1:30 p.m. Conference Adjournment

### **Return to top**

## **Poster Session A**

### **A - 1 Cytoplasmic Expression of Disulfide Bonded Proteins in E. Coli**

Paul H. Bessette and George Georgiou

Department of Chemical Engineering, University of Texas at Austin, USA

Xiaoming Zhang

Department of Molecular Biology, University of Texas at Austin, USA

### **A - 2 Secretory Production of Recombinant Proteins Using an Artificial Signal Sequence**

Sang Yup Lee and Jong Hyun Choi

Chemical Engineering Department, Korea Advanced Institute of Science and Technology, Korea

### **A - 3 High Level Production of Human Leptin and Its Purification**

Sang Yup Lee and Ki Jun Jeong

Department of Chemical Engineering, Korea Advanced Institute of Science and Technology, Korea

**A - 4 Amino Acid Depletion Model and Experimental Validation for Recombinant *E. coli***

Kevin J.-R. Clark, Fu'ad Haddadin and Sarah W. Harcum

Department of Chemical Engineering, New Mexico State University, Las Cruces, NM

**A - 5 Characterisation of Plasmid DNA Production in *Escherichia Coli* Fermentation for Gene Therapy**

R.D. O'Kennedy and E. Keshavarz-Moore

Advanced Centre for Biochemical Engineering, Department of Biochemical Engineering, University College London, UK

**A - 6 Extremophilic Enzymes: Heterologous Gene Expression and Use as Biocatalysts**

Dr. Julian B. Chaudhuri, Dr. Helen Connaris, Gerald Sellek, Prof. Michael Danson and Dr. David Hough

Department of Chemical Engineering, University of Bath, UK

**A - 7 Production of Antifungal Molecules Using Hydrothermal Marine Bacteria**

Joseph Boudrant

Laboratory of Chemical Engineering Sciences-CNRS, France

Joël Coulon and Roger Bonaly

UMR UHP-CNRS, Biochimie Microbienne, Faculté de Pharmacie, France

Georges Barbier

Laboratory of Microbiology, IFREMER, Centre de Brest, France

Jacques Dietrich

Laboratory of Biotechnology, IFREMER, Centre de Brest, France

**A - 8 Extracellular Proteinases from Psychrotrophs: Keratinase**

Quamrul Hasan and Daniel G. Moran

Procter & Gamble Far East, Inc., Research & Development Division, Japan

Yasutaka Morita and Eiichi Tamiya

Japan Advanced Institute of Science and Technology, The School of Materials Science, Japan

**A - 9 Study of Dynamic Response to Growth Rate Oscillation in Recombinant *Bacillus Subtilis* Culture**

S. Chuen-Im

Biotechnology and Biochemical Engineering Group, Department of Food Science and Technology, University of Reading, UK

D.L. Pyle and H.C. Lynch

Department of Food Science & Technology, University of Reading, UK

**A - 10 Modelling of Baeyer-Villiger Monooxygenase Biocatalytic Reactions**

Matthew Hogan and John Woodley

Department of Biochemical Engineering, University College London, UK

**A - 11 Axial Flow Up-Pumping Impellers in Industrial Fermentation Processes**

Samun K. Dahod, Ph.D.

Abbott Laboratories, USA

**A - 12 *Is the Production of Biomaterials by Fermentation Economically Viable?***Anton P.J. Middleberg

Department of Chemical Engineering, University of Cambridge, UK

Yao Ling and Richard J. van Wegen

University of Adelaide, Australia

**A - 13 *Mathematical Model of Inorganic Phosphate Controlled Expression of Dengue Envelope Protein by *Saccharomyces Cerevisiae****Anan Tongta

School of Bioresources and Technology, King Mongkut's University of Technology Thonburi, Thailand

Chulee Yompakdee

Pilot Plant Development and Training Institute, King Mongkut's University of Technology Thonburi, Thailand

**A - 14 *Production and Recovery of Human Endostatin from *Pichia Pastoris****Joseph Shiloach, Loc Trinh and Santosh Noronha

Biotechnology Unit, LCDB, NIDDK, National Institute of Health, USA

**A - 15 *NADH Oscillations in Brewers Yeast***Claus Emborg

Center for Process Biotechnology, Department of Biotechnology, Technical University of Denmark, Denmark

Jes Tobiassen

Alfred Jørgensen Lab, Copenhagen

**A - 16 *Exploiting the Diversity in Cell Wall Hydrolases of *Trichoderma Reesei****

H.J. Meerman

Genencor International, Inc.

**A - 17 *Cultivation and Direct Regeneration System of Embryogenic Rice Callus Using Macroporous Support***Hiroyuki Honda, Kwan Hoon Moon, Chunzhao Liu and Takeshi Kobayashi

Department of Biotechnology, Graduate School of Engineering, Nagoya University, Japan

**A - 18 *Glycosylational Engineering in Insect Cells: Effects of Gal T Overexpression***Michael J. Betenbaugh, Eric Ailor and Y.C. Lee

Department of Chemical Engineering, Johns Hopkins University, USA

Noriko Takahashi

Nakano Vinegar Co. Ltd.

Donald Jarvis

University of Wyoming, USA

**A - 19 *Manipulation of the Apoptosis Pathway in Mammalian Cell Culture***Tina M. Sauerwald, Bruno Figueroa, Jr., J. Marie Hardwick and Michael J. Betenbaugh

Department of Chemical Engineering, Johns Hopkins University, USA

**A - 20 *Impact of Dolichol Monophosphate Supplementation on Recombinant Gamma-Interferon Glycosylation***Inn H. Yuk and Prof. Daniel I.C. Wang

Massachusetts Institute of Technology, USA

**A - 21 Strategies for Production of High Titer Herpes-Based Gene Therapy Vectors**

J.C. Wetchuck, A. Ozuer, B. Russell and M.M. Ataai  
Chemical Engineering, University of Pittsburgh, USA

J. Glorioso

Molecular Biology and Biochemistry, University of Pittsburgh, USA

**A - 22 High Density Culture of Panax Notoginseng Cells in Bioreactors for Saponin Production**

H. Yao and J.J. Zhong

State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, China

**A - 23 Phase Formation Behavior and Kinetics of the Separation of Phases and Proteins in Aqueous Two-Phase Systems**

B. A. Andrews, E. Huenupi, R. Muñoz, and J. A. Asenjo

Center for Biochemical Engineering and Biotechnology, Department of Chemical Engineering, University of Chile, Santiago, Chile

**A - 24 Effects of Bone Marrow Architectures on Oxygen Tensions and Gradients Experienced by Hematopoietic Cells in vivo**

Dominic Chow

Northwestern University

**A - 25 Molecular Analysis of *spo0A* in *Clostridium Acetobutylicum* - Progress Towards Decoupling Stationary Phase Phenomena in Solventogenic Bacteria**

Latonia M. Harris

Northwestern University

**A - 26 Surface Display of Enzymes as a Selective Screening System**

Jac-Gu Pan

Korea Research Institute of Bioscience and Biotechnology

[Return to top](#)

**Poster SessionB**

**B - 1 Engineering Microorganisms for Heavy Metal Removal**

Douglas S. Clark, J.D. Keasling, Clifford L. Wang, Sang-Weon Bang and Andrew C. Magyarosy  
Department of Chemical Engineering, University of California-Berkeley, USA

**B - 2 Application of Flow Cytometry to Study Substrate and Product Toxicity in the Indene Bioconversion**

Ashraf Amanullah

Merck and Co. / University College London, Department of Biochemical Engineering, UK

C.J. Hewitt, A.W. Nienow

School of Chemical Engineering, The University of Birmingham, UK

M. Chartrain, B.C. Buckland and S.W. Drew

Department of Bioprocess Research and Development, Merck Research Laboratories, Merck and Co. Inc., USA



J.M. Woodley  
Department of Biochemical Engineering, University College London, UK

**B - 3 Sedimentation of Manganese/Iron Ore Fines from Wash Water of an Ore Washing Plant**  
A.D. Agate  
Agharkar Research Institute

**B - 4 DNA Cleavage and Biological Activities of Substituted Naphthyle Imides**  
Dong-Zhi Wei, Dong-Hui Zhu and Jiang-Chao Qian  
Research Institute of Biochemistry, State Key Laboratory of Bioreactor Engineering, East China University of Science & Technology, China  
Tian-Bao Huang and Xu-Hong Qian  
Institute of Pharmaceuticals & Pesticides, East China University of Science & Technology

**B - 5 Bioconversion in a Membrane Bioreactor-Experimental and Computer Simulated Comparisons**  
Preeta Tyagi  
Youngsoft, Inc.  
Prof. S.N. Upadhyay  
Institute of Technology, India

**B - 6 A Rapid and Rational Method for Selection of Cofactor Requiring Processes**  
Katie C. Thomas and John M. Woodley  
The Advanced Centre for Biochemical Engineering, Department of Biochemical Engineering, University College London, UK

**B - 7 Directed Discovery and Analysis of Antisense Oligonucleotides**  
Martin L. Yarmush, Charles M. Roth, S. Patrick Walton, and Arul Jayaraman  
Center for Engineering in Medicine, Massachusetts General Hospital, USA

**B - 8 Design and Discovery of Poly-(1-3)-Trans-(2-2)-GPDD from *Acholeplasma Laidlawii***  
L.L. Matz  
Matz & Associates, USA

**B - 9 Thermostable Peptide Ligase in DMF**  
Liujin Zhu and Wu Yukie  
Institute of Biophysics, Chinese Academy of Sciences, China  
Yang Yonghua and Yang Shengli  
Shanghai Research Center of Biotechnology, CAS, China

**B - 10 At Play in the Fields of Molecular Biology: Antibody Affinity Maturation via Combinatorial Genetics**  
Jennifer Anne Maynard, Brent L. Iverson and George Georgiou  
Department of Chemical Engineering, University of Texas at Austin

**B - 11 Protein Engineering of a Lytic *b*-1,3-Glucanase Enzyme able to Permabilize the Yeast Cell Wall**  
B.A. Andrews, A. Olivera, M. Casas, O. Salazar, J. Molitor and J.A. Asenjo  
Center for Chemical Engineering, Department of Chemical Engineering, University of Chile, Chile

**B - 12 Aqueous Two-Phase Processes for the Recovery of Aroma Compounds Produced by Micelial**

**Cultures**

Marco Rito-Palomares and Alejandro Negrete  
Centro de Biotecnología-ITESM, Mexico  
Leobardo Serrano and Enrique Galindo  
Instituto de Biotecnología-UNAM, Mexico

**B - 13 Peptide Affinity Chromatography of Human Clotting Factor VIII: Column Experiments with Peptides Derived From the Factor VIII-Binding von Willebrand Factor Region**

Karin Amatschek, Rainer Hahn, Roman Necina and Alois Jungbauer  
Institute for Applied Microbiology, Austria  
Djuro Josic  
Octapharma, Vienna

**B - 14 Selection of Optimum Affinity Tags for Immobilized Metal Affinity Chromatography Using a Phage Display Peptide Library**

Richard Pasquinelli, Ya Chen, Richard R. Koepsel, Rex Shepherd and Mohammad Ataai  
Department of Chemical Engineering & Center for Biotechnology & Bioengineering, University of Pittsburgh, USA

**B - 15 Human Blood Clotting Factor VIII-Binding Peptides Derived from a Combinatorial Peptide Library: Screening, Characterization and Application for Affinity Chromatography of Recombinant and Plasma Derived FVIII**

Alois Jungbauer, Karin Amatschek, Roman Necina and Eva Schallaun  
Institute of Applied Microbiology, University of Agriculture,  
Forestry & Biotechnology, Austria  
Horst Schwinn and Djuro Josic  
Octapharma Pharmazeutika Produktionsges.m.b.H., Austria

**B - 16 Microscale Gradients of Biomolecules for Cell Adhesion and Neurite Extension**

D.N. Adams, C.L. Hypolite, E.Y. Kao and W.-S. Hu  
Department of Chemical Engineering and Materials Science, University of Minnesota, USA  
M.D. Distefano  
Department of Chemistry, University of Minnesota, USA  
P.C. Letourneau  
Department of Cell Biology and Neuroanatomy, University of Minnesota, USA

**B - 17 Development of a System for Advanced Life Support in Space Based on a Loop of Bioreactors**

F. Gòdia, J. Albiol, J.L. Montesinos, A. Vernerey, J. Pérez, F. Cabello, N. Creus, A. Morist and X. Mengual  
Departament d'Enginyeria Química, Universitat Autònoma de Barcelona, Spain  
C. Lasseur  
European Space Research and Technology Center, European Space Agency, The Netherlands

**B - 18 Cysteine Mutations and the Cellular Redox Environment are Critical for the in vivo Folding and Assembly of the Non-Disulfide Bonded P22 Tailspike Protein**

Anne Skaja Robinson  
University of Delaware

**B - 19 Effect of Arginine on the Oxidative Renaturation of Hen Egg-White Lysozyme**

Lars Waldmann, Rainer Rudolph and Eliana De Bernardez Clark  
Department of Chemical Engineering, Tufts University and the Institute for Biotechnology, University  
of Halle, Germany

**B - 20 *Antemortem Diagnosis of Transmissible Spongiform Encephalopathy***

Kelvin H. Lee  
Cornell University

**B - 21 *Screening of Large Cell Populations for 'Enzyme Evolution'***

Amihay Freeman  
Tel Aviv University

**[Return to top](#)**